



5-2011

Novel Insights into Ubiquitin-like Protein E1-E2 Interactions

Asad Taherbhoy

University of Tennessee Health Science Center

Follow this and additional works at: <https://dc.uthsc.edu/dissertations>



Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Medical Biochemistry Commons](#), and the [Medical Cell Biology Commons](#)

Recommended Citation

Taherbhoy, Asad , "Novel Insights into Ubiquitin-like Protein E1-E2 Interactions" (2011). *Theses and Dissertations (ETD)*. Paper 261.
<http://dx.doi.org/10.21007/etd.cghs.2011.0310>.

This Dissertation is brought to you for free and open access by the College of Graduate Health Sciences at UTHSC Digital Commons. It has been accepted for inclusion in Theses and Dissertations (ETD) by an authorized administrator of UTHSC Digital Commons. For more information, please contact jwelch30@uthsc.edu.

Novel Insights into Ubiquitin-like Protein E1-E2 Interactions

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Biomedical Sciences

Track

Cell Biology and Biochemistry

Research Advisor

Brenda A. Schulman, Ph.D.

Committee

Linda M. Hendershot, Ph.D. Lawrence M. Pfeffer, Ph.D. Stephen White, Ph.D. Jie Zheng, Ph.D.

DOI

10.21007/etd.cghs.2011.0310

Comments

One year embargo expired May 2012

NOVEL INSIGHTS INTO UBIQUITIN-LIKE PROTEIN E1-E2 INTERACTIONS

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Asad Taherbhoy
May 2011

Chapter 2 © 2010 Wang et al.
All other material © 2011 by Asad Taherbhoy.

DEDICATION

To my wife for her constant love and support,
and my daughter for brightening each day!
To my parents, grandparents and in-laws
for all they have done over the years.

ACKNOWLEDGEMENTS

Above all, I would like to thank my mentor, Brenda Schulman, for immersing me into the world of cutting-edge research and teaching me what it takes to keep doing such great science. Your training will serve me well always. I also want to thank Brenda for creating a perfect environment for doing science – both in terms of recruiting fantastic people and providing great resources. I want to thank my committee members: Linda Hendershot, Lawrence Pfeffer, Stephen White and Jie Zheng for the time and effort they put into guiding me through graduate school.

I would like to thank the members (past and present) of the Schulman lab for all that they have taught me and making this journey a fun one: Danny Huang who taught me much about science, especially biochemistry; Hari Kamadurai for all the great scientific (and non-scientific) discussions over the years; Danny Scott for his insightful scientific suggestions and ITC training; Dave Duda for all the wonderful crystallography tips and tricks; Billy Dye for the guidance with cloning; David Miller for keeping the lab fully stocked and equipment always running; Jing Wang and Allison Williams for their help on projects we did together; Steve Kaiser for all his insights, discussions and proofreading; Mathew “Matt-man” Calabrese for taking the time to teach; Min Zhuang, Judith Souphron and Shanshan Yu, fellow graduate students, for sharing parts of the same journey; Shelia Boezman for all the care and administrative support (not forgetting the chocolates, candies and doughnuts); Julie Monda, for spreading the joy (and the homemade cookies); Steven Seyedin for all the fun times; Woddy Hunt and Allan “DDR” Deng for being ever helpful; Laura Borg for introducing me to the Schulman Lab protein preps, Jeremy Frye for the dissertation tips, Saugata Hazara, Catherine Regni, Zeynep Tocgoz, Jenny Olszewski and Allison Voss for being great lab mates.

I would also like to thank the following people their tremendous help along the way: Steven Tait and Doug Green for all their ideas and experiments on the in vivo autophagy work; Amanda Nourse for the analytical ultracentrifugation data; Darcie Miller for help with data collection and crystallography; John Bollinger and John Rogers for the computer/IT support.

Finally, a big THANK YOU to my family and friends that have made this possible!

ABSTRACT

Posttranslational modification of macromolecules by ubiquitin-like proteins (UBLs) such as ubiquitin, Sumo and NEDD8 regulate a vast array of processes in the cell. Transfer of UBLs to their target generally occurs by a series of molecular handoffs down an E1-E2-E3 cascade. We are interested in understanding how E1-E2 pairs interact and mediate UBL transfer. To this effect, we studied two E1-E2 pairs: the Sumo pathway (Sumo utilizes a canonical E1 and E2) and the Atg8 pathway (Atg8 is a UBL involved in autophagy that utilizes a non-canonical E1-E2 pair).

Sumo conjugation is initiated by the heterodimeric Aos1-Uba2 E1 enzyme (in humans called Sae1-Uba2), which activates Sumo's C-terminus, binds the dedicated E2 enzyme Ubc9, and promotes Sumo C-terminal transfer between the Uba2 and Ubc9 catalytic cysteines. In this study we determined crystal structures of the C-terminal ubiquitin-fold domain (ufd) from yeast Uba2 (Uba2^{ufd}), alone and in complex with Ubc9. The overall superimposition of Uba2^{ufd} and Ubc9 between yeast and human suggests conservation of fundamental features of Sumo conjugation. Docking the Uba2^{ufd}-Ubc9 and prior full-length human Uba2 structures allows generation of models for steps in Sumo transfer from Uba2 to Ubc9, and supports the notion that Uba2 undergoes remarkable conformational changes during the reaction. Comparisons to previous structures from the Nedd8 cascade demonstrate that UBL cascades generally utilize some parallel E1-E2 interaction surfaces. In addition, the structure of the Uba2^{ufd}-Ubc9 complex reveals interactions unique to Sumo E1 and E2. Comparison with a previous Ubc9-E3 complex structure demonstrates overlap between Uba2 and E3 binding sites on Ubc9, indicating that loading with Sumo and E3-catalyzed transfer to substrates are strictly separate steps. The results suggest mechanisms establishing specificity and order in Sumo conjugation cascades.

Atg7 is a noncanonical, homodimeric E1 enzyme that interacts with the noncanonical E2 enzyme, Atg3, to mediate conjugation of the ubiquitin-like protein (UBL) Atg8 during autophagy. Here we report that the unique N-terminal domain of Atg7 (Atg7^{NTD}) recruits a unique “flexible region” from Atg3 (Atg3^{FR}). The structure of an Atg7^{NTD}-Atg3^{FR} complex reveals hydrophobic residues from Atg3 engaging a conserved groove in Atg7, which mutational analyses show is important for Atg8 conjugation. The structure, along with homology of the Atg7 C-terminal domain (Atg7^{CTD}) to canonical E1s and their bacterial antecedents, allows modeling of a full-length, dimeric Atg7~Atg8-Atg3 complex. The model and supporting biochemical data provide a rationale for Atg7 dimerization: Atg8 is transferred in trans from the catalytic Cys of one Atg7 protomer to Atg3 bound to the N-terminal domain of the opposite Atg7 protomer within the homodimer. The studies reveal a novel E1~UBL-E2 architecture for enzymes mediating autophagy.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	1
UBLs	1
An Overview of UBL Transfer Down the E1-E2-E3 Cascade	3
Prokaryotic Ancestors of E1s	3
The Sumo Pathway	5
Sumo	5
Sumo E1 – the Bridge between Sumo and Its E2	6
Sumo Conjugation to Substrates	11
UBL Systems in the Autophagy Pathway	11
Autophagy and Its Two Dedicated UBL Pathways	11
Atg8 – a UBL with a Unique Target and Function	13
Atg7 and Atg3 – Atg8’s Non-canonical E1 and E2	14
CHAPTER 2. CRYSTAL STRUCTURE OF UBA2^{UFD}-UBC9: INSIGHTS INTO E1-E2 INTERACTIONS IN SUMO PATHWAYS	17
Introduction	17
Results and Discussion	18
Structure of an Isolated Ubiquitin-fold Domain from Uba2 (Uba2 ^{ufd})	18
Structure of a Uba2 ^{ufd} -Ubc9 Complex	25
Details of Uba2 ^{ufd} -Ubc9 Interactions	25
Conservation of Ubc9 and Uba2 Interacting Residues Across Species	27
Comparison of Uba2 ^{ufd} -Ubc9 with Prior E1 ^{ufd} Structures from Other Ubiquitin- like Protein Pathways	30
Docking Uba2 ^{ufd} -Ubc9 onto Prior Full-length Human Aos1-Uba2 Structures	30
Implications for E1-E2-E3 Cascades in the Sumo Pathway	33
Conclusions	35
Materials and Methods	36
Cloning, Protein Expression and Purification	36
Crystallization, Data Collection, and Structure Determination and Refinement	37
Assays for Formation of a Ubc9~Sumo Thioester Conjugate	38
CHAPTER 3. STRUCTURE OF AN ATG7-ATG3 SUB-COMPLEX REVEALS A DISTINCTIVE INTERACTION AND MECHANISM FOR TRANSTHIOLATION UNIQUE TO THE AUTOPHAGY PATHWAY	39
Introduction	39
Results and Discussion	43
Atg7 N-terminal Domain Recruits the Atg3 “Flexible Region”	43
Structural Basis for Atg7 ^{NTD} -Atg3 ^{FRpep} Interactions	46
Role for Atg7 ^{NTD} -Atg3 ^{FRpep} Interface in Atg7-Atg3 Interactions and in the Transfer of Atg8	54
The Atg7 ^{NTD} Binding Surface for Atg3 ^{FRpep} is Conserved Across Species	58
Model for Transfer of Atg8 from Atg7 onto Atg3: a Unique Trans Mechanism Requiring Optimal Spacing between the Atg7 ^{NTD} and Atg7 ^{CTD}	63
Implications for E1~UBL-E2 Cascades in Autophagy	67

Materials and Methods.....	68
Constructs, Protein Expression and Purification	68
Analytical Ultracentrifugation (AUC)	69
Isothermal Titration Calorimetry (ITC)	71
Purification, Crystallization and Structure Determination for Atg7 ^{NTD}	72
Purification, Crystallization and Structure Determination for Atg7 ^{NTD} -Atg3 ^{FR_{pep}} ...	73
Purification, Crystallization and Structure Determination for Atg7 ^{NTD} (P283D).....	73
Pull-down Assays.....	73
Generation of Model Showing Transfer in Trans	74
Pulse-chase Transfer of [³² P]Atg8 from Atg7 to Atg3	74
Cell-based Assays	77
CHAPTER 4. IMPLICATIONS AND FUTURE DIRECTIONS	78
Expansion on the E1 Theme	78
The Adenylation Domain and Beyond	78
Flexible Tethering of Domains for Functionality	81
The Need for Homodimerization	82
The Remaining Mystery of Atg7 Function	83
Therapeutic Targeting of the Autophagy Pathway	83
The Yin and Yang of Autophagy	83
Specific Targeting of Autophagy Proteins with Small Molecules.....	85
LIST OF REFERENCES	90
VITA	103

LIST OF TABLES

Table 2-1.	Crystallographic and refinement statistics for Uba2 ^{ufd} and Uba2 ^{ufd} -Ubc9.	19
Table 3-1.	Crystallographic and refinement statistics for Atg7 ^{NTD} , Atg7 ^{NTD} -Atg3 ^{FR_{pep}} and Atg7 ^{NTD} (P283D).	52

LIST OF FIGURES

Figure 1-1.	Transfer of a UBL down the E1-E2-E3 cascade.....	2
Figure 1-2.	Structural alignment of Ubiquitin-like proteins.	3
Figure 1-3.	The Adenylation domain is at the core of all E1s.	4
Figure 1-4.	The structural domains of Sumo's E1 and the mechanism of canonical E1 thiolester formation.....	7
Figure 1-5.	E2s from the Sumo, ubiquitin and Nedd8 pathways share structural similarity.....	9
Figure 1-6.	Structural studies on the E1 for the Nedd8 pathway reveals the mechanism for the transthiolester reaction.	10
Figure 1-7.	Overview of autophagy and its two UBL systems.....	12
Figure 1-8.	Comparison of Atg3 with Ubc9.....	15
Figure 2-1.	Representative electron density.....	21
Figure 2-2.	Overall structure of the Uba2 ^{ufd} -Ubc9 complex.....	22
Figure 2-3.	Ubc9 binds W-shaped groove in Uba2 ^{ufd}	24
Figure 2-4.	Details of the Uba2 ^{ufd} -Ubc9 interface.....	26
Figure 2-5.	Effects of mutations in either yeast Uba2 or Ubc9 on forming the thiolester-linked Ubc9~Sumo thiolester conjugate.	28
Figure 2-6.	Specificity of E1 ^{ufd} -E2 interactions.....	29
Figure 2-7.	Sequence comparisons of Ubc9, Uba2 ^{ufd} , and corresponding regions of E2s and E1s for other UBLs.....	31
Figure 2-8.	Implications of Uba2 ^{ufd} -Ubc9 structure for Smt3 transfer from Uba2 to Ubc9.	32
Figure 2-9.	Implications of Uba2 ^{ufd} -Ubc9 structure for Sumo cascades.....	34
Figure 3-1.	Structures of UBLs and their corresponding adenylation activation domains.	40
Figure 3-2.	Atg7 ^{CTD} multiple sequence alignment.	41

Figure 3-3.	Monomeric adenylation domains bound to their corresponding UBL and a molecule of ATP.....	42
Figure 3-4.	Atg7 ^{NTD} multiple sequence alignment.	44
Figure 3-5.	Atg7 is a homodimer, with the Atg7 ^{CTD} mediating the dimerization.	45
Figure 3-6.	Mapping the high-affinity binding site on Atg7 and Atg3 that mediates their interaction.	47
Figure 3-7.	Mapping the minimal Atg3 ^{FR} region required for binding the Atg7 ^{NTD}	48
Figure 3-8.	Structural features of Atg7 ^{NTD}	49
Figure 3-9.	Final 2F _o -F _c electron density maps of Atg7 ^{NTD}	50
Figure 3-10.	Structure of the Atg7 ^{NTD} -Atg3 ^{FRpep} complex.	51
Figure 3-11.	The FR region of Atg3 is flexible.	55
Figure 3-12.	The Atg3 ^{FRpep} mediates interactions with Atg7 ^{NTD} via a hydrophobic interface.....	56
Figure 3-13.	Atg7 ^{NTD} -Atg3 ^{FRpep} interface is important for Atg7-Atg3 interactions and for the transfer of Atg8.	57
Figure 3-14.	Overlay of the crystal structure of Atg7 ^{NTD} with that of Atg7 ^{NTD} (P283D).....	59
Figure 3-15.	Atg3 ^{FR} binds a conserved surface on Atg7 ^{NTD}	60
Figure 3-16.	Mouse Atg7 uses the same conserved surface as its yeast counterpart to bind mouse Atg3 ^{FR}	61
Figure 3-17.	Mutation on Atg7's Atg3 ^{FR} binding surface leads to defects in autophagy.	62
Figure 3-18.	Transthiolation in trans: a working model for transfer of Atg8 from Atg7 to Atg3.....	64
Figure 3-19.	Transthiolation model of dimeric Atg7 bound to two molecules of Atg3, Atg8 and ATP.....	65
Figure 3-20.	Atg7 performs the transthiolation reaction in trans and has a requirement for its linker length between the Atg7 ^{NTD} and Atg7 ^{CTD}	66
Figure 3-21.	Gels showing Atg7 ^{trans} and Atg7 ^{cis} during purification.	67
Figure 4-1.	MccB is a prokaryotic protein with an adenylation domain at its core.....	79

Figure 4-2.	Beyond the MoeB/ThiF-like adenylation domain in eukaryotic E1s.....	80
Figure 4-3.	MLN4924 forms a covalent adduct with the C-terminus of Nedd8.....	86
Figure 4-4.	Mechanism by which MLN4924 inhibits the E1 for Nedd8.....	88
Figure 4-5.	Comparison of the MDM2-p53 ^{pep} structure with the Atg7 ^{NTD} -Atg3 ^{FR_{pep}} structure.	89

LIST OF ABBREVIATIONS

A β	β -amyloid
AMP	Adenosine 5'-monophosphate
Ape1	Aminopeptidase I
Atg	autophagy related
Atg3 ^{FR}	Atg3 Flexible Region
Atg3 ^{FRpep}	synthetic peptide encompassing Atg3 residues 128-144
Atg3 ^{HR}	Atg3 Handle Region
Atg7 ^{C-ter}	Atg7 residues 572-630
Atg7 ^{CTD}	Atg7 C-terminal Domain
Atg7 ^{NTD}	Atg7 N-terminal Domain
ATP	Adenosine-5'-triphosphate
AUC	analytical ultracentrifugation
DUBs	deubiquitinating enzymes
E1	activating enzyme
E2	conjugating enzyme
E3	ligase
FPLC	fast protein liquid chromatography
ITC	isothermal titration calorimetry
MccC7	Microcin C7
MEFs	murine embryonic fibroblasts
MPN	Mpr1, Pad1 N-terminal
NEDD8	neural precursor cell-expressed developmentally down-regulated 8
PE	Phosphatidylethanolamine
RHD	rhodanase homology domain
rmsd	root mean square deviation
SIM	Sumo interacting motif
Smt3	Suppressor of migration inhibition factor 2
Sumo	Small ubiquitin-like modifier
UBL	ubiquitin-like protein
UCF1	UFM1-conjugating enzyme
ufd or E1 ^{ufd}	ubiquitin-fold domain
UFL1	UFM1-ligase
UFM1	ubiquitin-fold modifier 1
Urm1	ubiquitin-related modifier 1
WT	wild-type

CHAPTER 1. INTRODUCTION

Post-translational modification is a ubiquitously observed phenomenon in cells that alters the fate of the protein. Ubiquitin and ubiquitin-like proteins (referred to collectively as UBLs) are eukaryotic specific protein modifiers that get post-translationally attached to protein, via an E1-E2-E3 multi-enzyme cascade, to alter the fate of their target (Figure 1-1A). Compared to post-translational modification by small molecules such as phosphate or methyl groups, UBLs have a larger surface with increased chemical diversity for enhanced regulation of protein function [1]. Overall, post-translational modification by UBLs control numerous processes, including cell-cycle regulation, immune responses, protein transport, quality control in the endoplasmic reticulum and embryonic development to name a few, making it important to understand the mechanism behind how UBLs are transferred to substrates [2, 3].

My dissertation work focuses on E1-E2 interaction and transthiolation (transfer of a UBL from E1 to E2) mechanism – with the aim of helping us better understand how UBLs are moved down their cascades. Thus the introduction will be more E1-E2-centric in nature. Additionally, the emphasis will be on two particular UBL systems, namely Sumo and ATG8, since these systems comprise all the work presented here.

UBLs

UBLs such as ubiquitin, Sumo (Small Ubiquitin-related Modifier) and Nedd8 (Neural Precursor Cell-expressed Developmentally Down-regulated protein 8) are structurally similar proteins that are composed of a globular core that adopts a β -grasp fold and a flexible C-terminal tail that ends in a Gly-Gly motif (Figures 1-2A-2C) [4-7]. It is this C-terminal tail that is activated (through an adenylation reaction discussed below) for eventual conjugation to the target molecule [8-10]. UBLs are typically conjugated to ϵ -amino group on lysine residues and the N-termini of proteins. Some UBLs, such as ubiquitin and Sumo, have the ability to form UBL chains, where the C-terminus on the incoming UBL is conjugated to a lysine on a UBL already attached to substrate [3, 7, 11].

Interestingly, UBLs appear to have evolved from proteins within prokaryotic biosynthetic pathways [3, 12, 13]. MoaD and ThiS, structural homologues of UBLs, are involved in molybdopterin and thiamin biosynthesis pathways respectively [12, 13] (Figures 1-2D and 2E). Like UBLs, MoaD and ThiS are adenylated at their C-termini. In contrast to eukaryotic UBLs however, they do not post-translationally modify proteins or other macromolecules. Instead, they are transiently modified, carrying sulfur atoms at their C-termini, and are involved in sulfur transfer in their respective pathways [14].

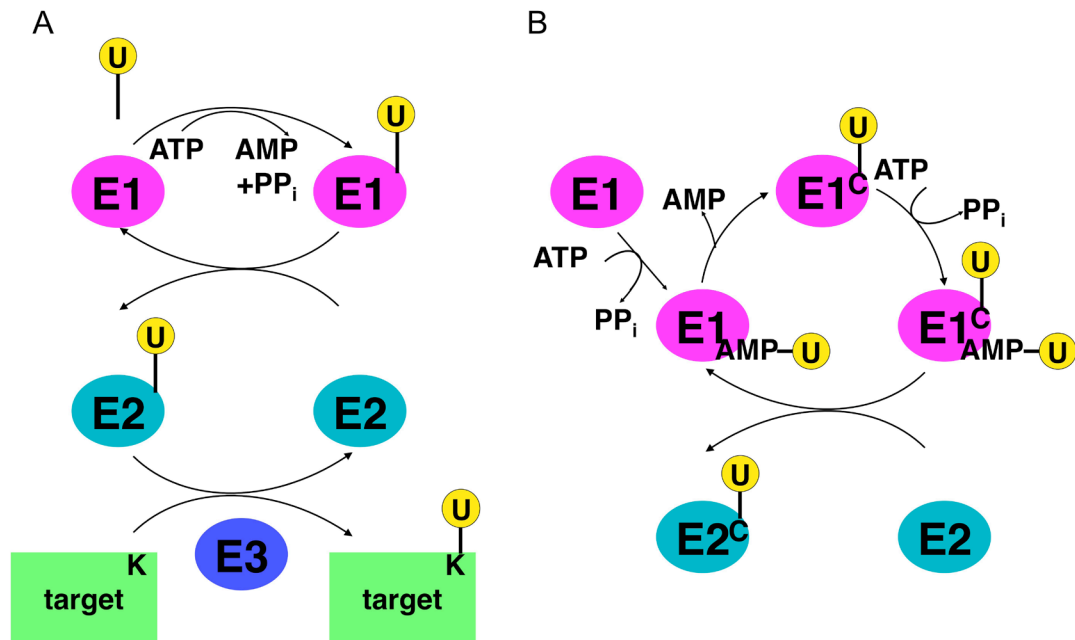


Figure 1-1. Transfer of a UBL down the E1-E2-E3 cascade.

(A) A UBL (yellow) is initially activated by an E1 (activating enzyme, magenta) while utilizing a molecule of ATP. The UBL then forms a thiolester linkage with the catalytic cysteine of the E1. Next, the E1 recruits an E2 (cyan) and transfers the UBL to the E2 catalytic cysteine. Finally, with the aid of an E3 (blue), the UBL is transferred to a lysine on the target macromolecule. (B) Canonical E1s have the ability to be double-loaded with two UBLs. Following thiolester formation between the catalytic cysteine of the E1 (denoted by a “C”) and UBL, the E1 can bind a second UBL in the adenylation site (denoted by a “AMP”).

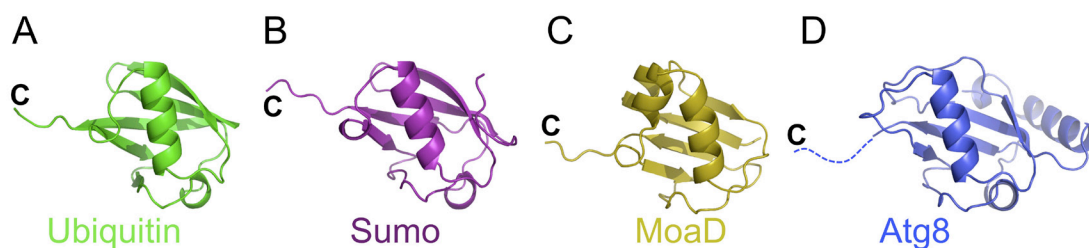


Figure 1-2. Structural alignment of Ubiquitin-like proteins.

(A) Ubiquitin (green, PDB code: 1UBQ). (B) Sumo (purple, PDB code: 1Y8R). (C) MoaD (gold, PDB code: 1JW9). (D) Atg8 (blue, PDB code: 2ZPN). C-terminal tail missing from the structure shown in dashed lines.

An Overview of UBL Transfer Down the E1-E2-E3 Cascade

For a UBL to be attached onto its target, it generally needs to proceed down a cascade of three enzymes: an E1 (activating enzyme), an E2 (conjugating enzyme), and an E3 (ligase) (Figure 1-1A). The ubiquitin system is the best understood UBL pathway and will therefore be used to describe this multi-enzyme cascade more elaborately. First, the E1 binds $\text{ATP} \cdot \text{Mg}^{2+}$ and ubiquitin, and catalyzes ubiquitin C-terminal acyl-adenylation [8-10]. This ubiquitin~adenylate forms a tight noncovalent complex with the E1 [10, 15] (the “~” represents a covalent bond, while “-” represents a non-covalent complex). This is followed by a nucleophilic attack by the E1 catalytic cysteine on the ubiquitin~adenylate to form an active E1~ubiquitin thioester complex [9, 10, 15]. Next, the E1 binds and adenylates a second ubiquitin molecule such that the E1 is now doubly loaded with ubiquitin – one bound covalently to the catalytic cysteine via a thioester and the second noncovalently as an adenylate [15, 16]. Finally, the E1 recruits an E2, the next enzyme in the cascade, for the transfer of ubiquitin (Figure 1-1B). This involves a transient noncovalent interaction between the E1 and E2 and a transthioesterification reaction, where the ubiquitin moiety is transferred from the E1’s catalytic cysteine to the E2’s [17, 18]. An E3 would then recruit the charged E2~ubiquitin and substrate for the final transfer of ubiquitin onto its target lysine residue [3, 14].

Prokaryotic Ancestors of E1s

Just as eukaryotic UBLs have their ancestry in prokaryotic biosynthetic pathways, so do the eukaryotic E1s that activate these UBLs. MoeB and ThiF are the enzymes that catalyze the C-terminal acyl-adenylation reactions of MoaD and ThiS, respectively (Figure 1-3A) [12, 13]. These proteins share sequence/structural homology with the

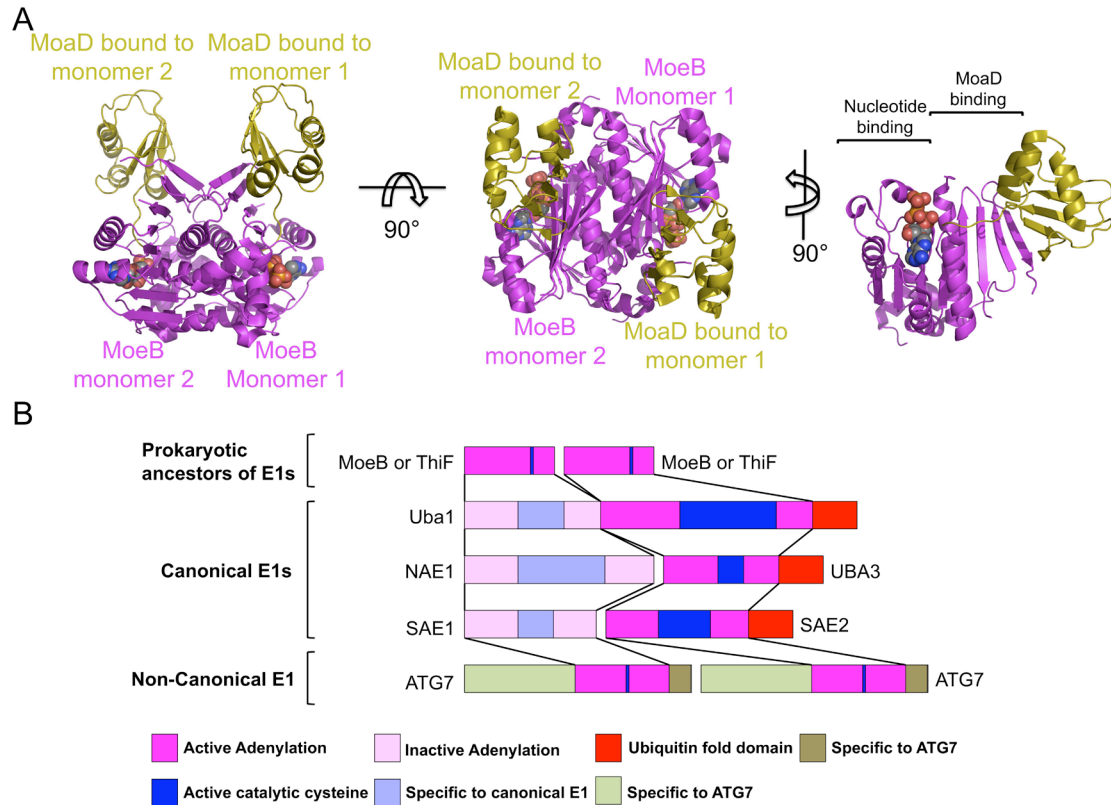


Figure 1-3. The Adenylation domain is at the core of all E1s.

(A) MoeB forms a homodimeric adenylation domain (magenta). Each monomer binds a molecule of MoaD (gold) and ATP (spheres), left and center panels. Right panel shows a MoeB monomer bound to MoaD. (B) A schematic representation showing the presence of MoeB/ThiF adenylation domains at the core of canonical and non-canonical E1s.

domain in eukaryotic E1s that bind UBLs and catalyze the adenylation reaction. Essentially, MoeB and ThiF are found at the core of all eukaryotic E1s (Figure 1-3B) [3, 12, 13].

Several crystal structures of MoeB and ThiF, with and without MoaD and ThiS, have provided great insight into how UBLs are recognized and activated [12, 13, 19]. MoeB and ThiF exist as homodimers with one active site per monomer (Figure 1-3) [12, 13]. The monomer can be further divided into two subdomains – an N-terminal subdomain and C-terminal subdomain (Figure 1-3A). The N-terminal subdomain is made up of a variant Rossmann nucleotide-binding fold that consists of a four-stranded parallel β -sheet surrounded by α -helices. This subdomain also contains the Gly-X-Gly-X-X-Gly nucleotide-binding motif, where ATP (Adenosine-5'-triphosphate) used in the adenylation reaction binds [12, 13]. The C-terminal subdomain is made up of a four-stranded anti-parallel β -sheet and adopts a fold distantly related to sugar binding proteins [12]. This C-terminal subdomain provides the bulk of the interaction with its prokaryotic UBL, the C-terminal tail of the UBL extending through this subdomain, ending in close proximity to the ATP [12, 13]. Another interesting feature of these homodimeric enzymes is the presence of an Arg finger. Essentially, an Arg residue from one monomer extends into the active site of the other monomer, binding the γ -phosphate of ATP, demonstrating the interdependence of the two monomers [12-14]. Lastly, each MoeB and ThiF monomer binds a Zn^{2+} at its C-terminus via tetrahedral coordination with four cysteines that exist as Cys-X-X-Cys motifs [12-14]. Overall, the features described here are present in eukaryotic E1s and sets the stage for understanding the core structure upon which all E1s are built. Addition of sequences at the N- or C-termini, or insertions within these primitive MoeB and ThiF enzymes give rise to modern E1s with functions specific for their partner UBLs.

E1s, in addition to their adenylation domains that are responsible for binding and activating UBL C-termini, have a range of other domains. E1s for ubiquitin (UBA1 and UBA6), Sumo (SAE1-UBA2) and NEDD8 (NAE1-UBA3) will be categorized as canonical E1s from here on based on their related domain structures and enzymatic mechanisms. In contrast, ATG7, the E1 for autophagy will be classified as a non-canonical E1. In general, canonical E1s are made up of a single polypeptide or are heterodimers, while non-canonical E1s are homodimeric [3] (Figure 1-3B). The Sumo E1 and ATG7 will be used as examples to discuss E1 structure, E1 mechanism and interaction with E2s for transthiolation, including specifics on the UBLs involved.

The Sumo Pathway

Sumo

In the process of Sumoylation, Sumo covalently modifies proteins, regulating a vast number of cellular processes such as transcription, nuclear transport and DNA damage [20]. Sumo shares only ~18% sequence identity with ubiquitin but has a virtually

identical three-dimensional structure (Figure 1-2B) [5]. *S. cerevisiae* expresses one Sumo polypeptide called Smt3 (suppressor of migration inhibition factor 2 (MIF-2)) while mammals express Sumo1-4 [7, 11]. It is uncertain however if Sumo4 is conjugated to protein in cells [21, 22]. Sumo2 and Sumo3 have 97% identity to one another (often being referred to as Sumo2/3) while having only ~50% sequence identity to Sumo1. Additionally, unlike Sumo2/3, Sumo1 lacks the ability to form polySumo chains. Many Sumoylation sites (lysine residues) lie in a Ψ -K-X-E/D motif, where Ψ is a large hydrophobic amino acid. Sumo2/3 (including *S. cerevisiae* Smt3) have this consensus motif at their N-termini whereas Sumo1 does not, accounting for the inability of Sumo1 to form polySumo chains [7].

Sumo E1 – the Bridge between Sumo and Its E2

E1s in general perform three functions: 1) they activate the C-terminal of the UBL, priming it for subsequent transfer down the cascade; 2) they have a catalytic cysteine that attacks the activated C-terminal tail, forming a thioester bond with the UBL – this allows for subsequent transfer of the UBL to the next enzyme in the pathway, the E2 and; 3) they recruit E2s and transfer the UBL to the catalytic cysteine of the E2 via a transthiolation reaction [23-25]. Just like other canonical E1s, Sumo E1 possesses three domains for each of the three functions mentioned above: an adenylation domain for activating Sumo's C-terminus; a catalytic cysteine domain that houses E1s catalytic cysteine and; a ubiquitin fold domain (ufd or E1^{ufd}), named for its structural resemblance to ubiquitin) for recruiting an E2 (Figure 1-4A) [24].

Canonical E1s have two MoeB/ThiF repeats making up their adenylation domains. The two repeats can be encoded in a single polypeptide (as N- and C-terminal halves) or in two separate subunits (Figure 1-3D). Notably, the adenylation domains of canonical E1s are pseudosymmetric in nature – having one active site per functional E1 – one MoeB/ThiF repeat binds to ATP•Mg²⁺ and UBL while the other provides structural stability and contributes the Arg finger that is crucial for catalysis [3, 23, 24, 26].

The canonical Sumo E1 is a heterodimer composed of SAE1 and UBA2 (Figure 1-4A). UBA2 primarily forms the three domains described above, which includes the active half of the pseudosymmetric adenylation domain. SAE1 folds into a single domain that dimerizes with the active adenylation domain of UBA2, giving rise to the pseudosymmetric adenylation domain with the Arg finger coming from SAE2 [24]. In addition to binding ATP•Mg²⁺, the adenylation domain of Sumo E1 binds a molecule of Sumo; specifically, Sumo1 is recognized by residues from UBA2 and no direct contacts are seen between the UBL and SAE1 [24]. It is within this adenylation domain that the Sumo E1 activates the C-terminus of the UBL, priming it for attack by the catalytic cysteine [24].

How Sumo (or other UBLs) got from the adenylation site onto the catalytic cysteine – a distance of ~30 Å remained a mystery until recently, when an ingenious

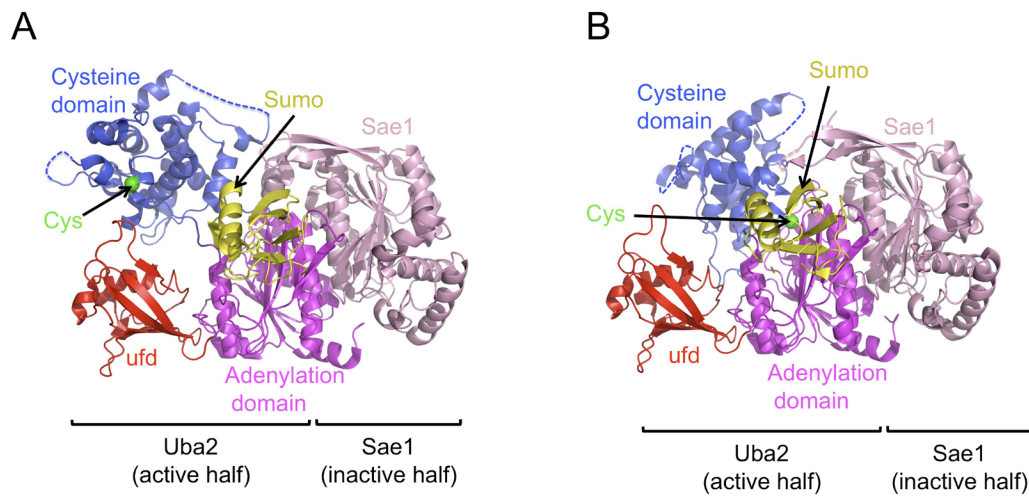


Figure 1-4. The structural domains of Sumo's E1 and the mechanism of canonical E1 thiolester formation.

(A) The E1 for Sumo is a heterodimer composed of Sae1 (pink) and Uba2 (PDB code: 3KYC). Uba2 possess three domain crucial to E1 function: 1) an adenylation domain (magenta) that binds and activates Sumo; 2) a cysteine domain (blue) that houses the catalytic cysteine (green) that attacks the adenylated Sumo moiety; and 3) a ubiquitin-fold domain (ufd, red) that binds the E2 for transthiolation. Sae1 is the inactive monomer in this complex, unable to bind ATP or Sumo. (B) The cysteine domain (blue) undergoes huge conformational change in order to reach the active site of adenylation, to attack the adenylated UBL for thiolester formation (PDB code: 3KYD).

chemical approach was taken to trap the E1 in a conformation that allowed visualization of the catalytic cysteine attacking the UBL C-terminus [27]. The structure revealed that the cysteine domain of the E1 rotates 130° and remodels itself to reach the UBL-adenylate in the adenylation domain (Figure 1-4B) [27]. Attack by the catalytic cysteine results in formation of a thioester bond between the catalytic cysteine and the C-terminal of Sumo [3, 27].

The E1 then needs to interact with its E2 for the transfer of Sumo onto the E2. Generally, E2s that function with the canonical E1s are composed of an elongated globular core that is ~150 amino acids and houses the catalytic cysteine which attacks the E1~UBL thioester to form an E2~UBL thioester [7]. Some E2s have an extension at the N- and/or C-terminus. Sumo proteins utilize a common E2 – Ubc9 – that is composed of only the globular core without any extensions on its termini (Figure 1-5A). Ubc9 is structurally highly similar to the ubiquitin and NEDD8 E2s (Figure 1-5B) [7].

How do Sumo E1 and Ubc9 interact to allow for the transfer of the UBL? The first insights into this came from a mutational study of *S. cerevisiae* Ubc9 that showed that the N-terminal α -helix and the loop between β -strands 1 and 2 (β 1- β 2 loop) (Figure 1-5A) were important for binding to the E1 [28]. This surface of Ubc9, in binding E1, was recently validated by NMR [29]. These features of the E1 binding sites are thought to be conserved among E2s for other UBL as two studies with E2s for ubiquitin revealed that mutations in the N-terminal α -helix decreased E2~ubiquitin thioester formation [30, 31]. Interestingly, the β 1- β 2 loop region was found to be longer than that found in E2s from other UBL pathways (Figure 1-5B) [32, 33]. Two studies published in 2005 on the Sumo and NEDD8 pathways demonstrated that the E1 ufd is responsible for binding the E2 [24, 25]. The Sumo E1 study showed biochemically the interaction between the E1^{ufd} and the E2, while the NEDD8 specific study solved a structure of an E1 ufd-E2 complex (Figure 1-6A). The structure revealed that the NEDD8 E1^{ufd}-E2 binding is mediated primarily by hydrophobic interaction, being further stabilized by hydrogen bonds and salt bridges [25]. However, superimposition of this complex onto the full-length E1 structure (by overlaying the E1^{ufd}s) revealed that the catalytic cysteine of the E2 faced away from that of the E1, being separated by >50 Å (Figure 1-6A) [25]. A landmark paper by Huang et al. in 2007 solved the structure of NEDD8's E1 in complex with two molecules of NEDD8 and NEDD8's E2, UBE2M [34]. This structure revealed a rotation of the ufd by ~120° about a hinge region that connects the ufd to the rest of the E1, now placing the two catalytic cysteines in close proximity, facing each other – allowing us to understand how the transthioylation reaction (transfer of UBL from E1 to E2) would occur in canonical E1s (Figure 1-6B). More recently, there has been a structure of NEDD8's second E2, UBE2F in complex with NEDD8 E1's ufd [35]. This interaction too is primarily hydrophobic in nature and similar to that between NEDD8 E1's ufd and UBE2M. The pliable nature of hydrophobic interactions allows for the accommodation of a large array of structures or sequences. There is however no structural information on how Ubc9 binds the Sumo E1^{ufd} (or that of ubiquitin E1^{ufd} with a ubiquitin E2). Such structural information will allow us to better understand how different UBL pathways specifically mediate E1-E2 interactions without any cross-talk.

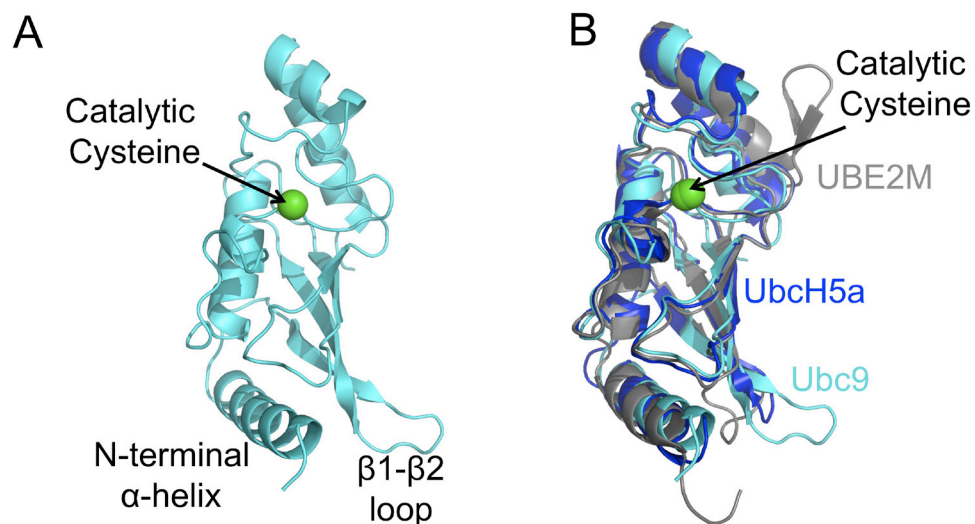


Figure 1-5. E2s from the Sumo, ubiquitin and Nedd8 pathways share structural similarity.

(A) Structure of Ubc9 (cyan, PDB code: 1U9A), the E2 for Sumo. Catalytic cysteine represented by a green sphere. (B) Overlay of Ubc9 with Ubc5a (blue, PDB code: 2C4P) and Ubc12 (grey, PDB code: 1Y8X), E2s from the ubiquitin and Nedd8 pathways respectively. Catalytic cysteine represented by a green sphere.

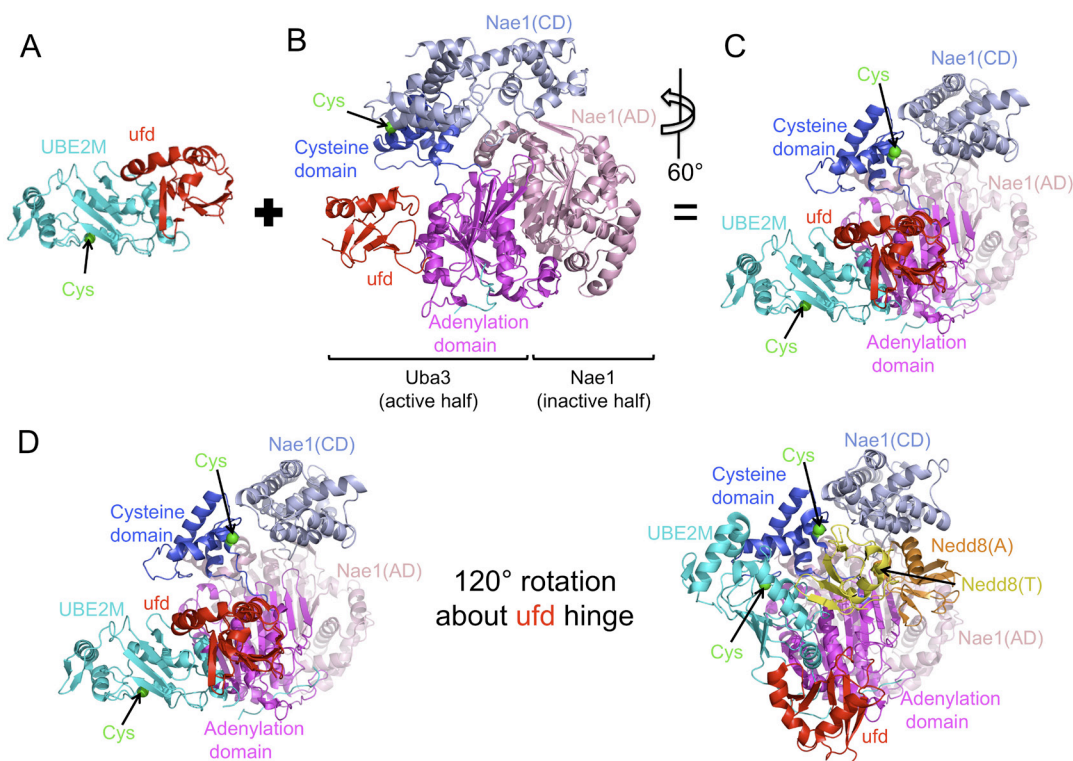


Figure 1-6. Structural studies on the E1 for the Nedd8 pathway reveals the mechanism for the transthiolation reaction.

(A) Structure of UBE2M (cyan) bound to Nedd8's E1^{ufl} (red). Catalytic cysteine shown as green sphere. PDB code: 1Y8X. (B) Structure of E1 for Nedd8 is composed of a heterodimer made up of Nae1 (AD : Adenylation Domain (pink, this is the inactive adenylation domain) and CD: Cysteine Domain (light blue)) and Uba3 (PDB code: 1TT5). Uba3 possess three domain crucial to E1 function: 1) an adenylation domain (magenta) that binds and activates Nedd8; 2) a cysteine domain (blue) that houses the catalytic cysteine (green) that attacks the adenylated Nedd8 molecule; and 3) a ubiquitin-fold domain (ufd, red) that binds the E2 for transthiolation. Nae1 is the inactive monomer in this complex, unable to bind ATP or Sumo. (C) Overlay of structures from parts (A) and (B) reveal a large gap between the two catalytic cysteines (green). These cysteines would need to approach each other in order to catalyze the transthiolation reaction. (D) Rotation about the ufd hinge allows the two catalytic cysteines to face each other within close proximity. Left panel from part (C), right panel from crystal structure with PDB code: 2NVU.

Sumo Conjugation to Substrates

The Sumo pathway is unique in the respect that Ubc9, the Sumo E2, can directly bind substrate and Sumoylate the target in the absence of an E3. Generally, proteins that get Sumoylated have the Ψ -K-X-E/D motif, with the lysine residue being modified by Sumo [11]. It is this motif that is recognized by Ubc9. Structural studies on the interaction between Ubc9 and substrate show that this motif exists in an extended conformation, with the Lys residue fitting into a hydrophobic groove on Ubc9. Residues flanking the Lys residue interact with Ubc9 via electrostatic interaction and hydrogen bonding – enabling recognition of this motif by Ubc9 [11, 36]. Although Ubc9 can directly bind and Sumoylate substrates, E3s do exist in the Sumo pathway that serve to enhance Sumo conjugation to targets and it is thought that majority of Sumoylation in cells requires E3s that are categorized as the PIAS proteins, Pc2 and RanBP2/Nup358 [7, 20].

An interesting question that remains to be answered is whether the Sumo E2 must dissociate from the E1 for binding to E3s. It has been shown for the ubiquitin and NEDD8 pathways that the E2 enzymes must dissociate from the E1 for subsequent interaction with the E3. Though current evidence points to the likelihood that the Sumo pathway may also require the E2 to dissociate from the E1 before engaging the E3, more work needs to be done towards understanding this.

UBL Systems in the Autophagy Pathway

Autophagy and Its Two Dedicated UBL Pathways

Macroautophagy (hereafter autophagy) refers to the lysosome-mediated degradation of macromolecules and organelles and is an evolutionarily conserved pathway in eukaryotes [37-39]. In autophagy, a double-membrane structure known as the autophagosome engulfs a portion of the cytoplasm and subsequently fuses with a lysosome (vacuole in yeast) where the cytosolic contents are degraded and recycled back to the cytoplasm for reuse (Figure 1-7A) [37, 40].

One of the best-known triggers of autophagy in yeast is nutrient deprivation [41]. However, autophagy is now known to have further roles in response to certain physiological demands, and is implicated in many diseases [42, 43]. Autophagy is required to remove non-functional organelles and maintain protein homeostasis to prevent liver, cardiac and metabolic diseases [44-50]; autophagy serves important roles in immunity and infection [43, 51, 52]; autophagy also suppresses tumor development and is implicated in many cancers [53-55]; defects in autophagy are associated with numerous neurodegenerative disorders [56-58]; and autophagy is also thought to contribute to many processes associated with aging [59-61].

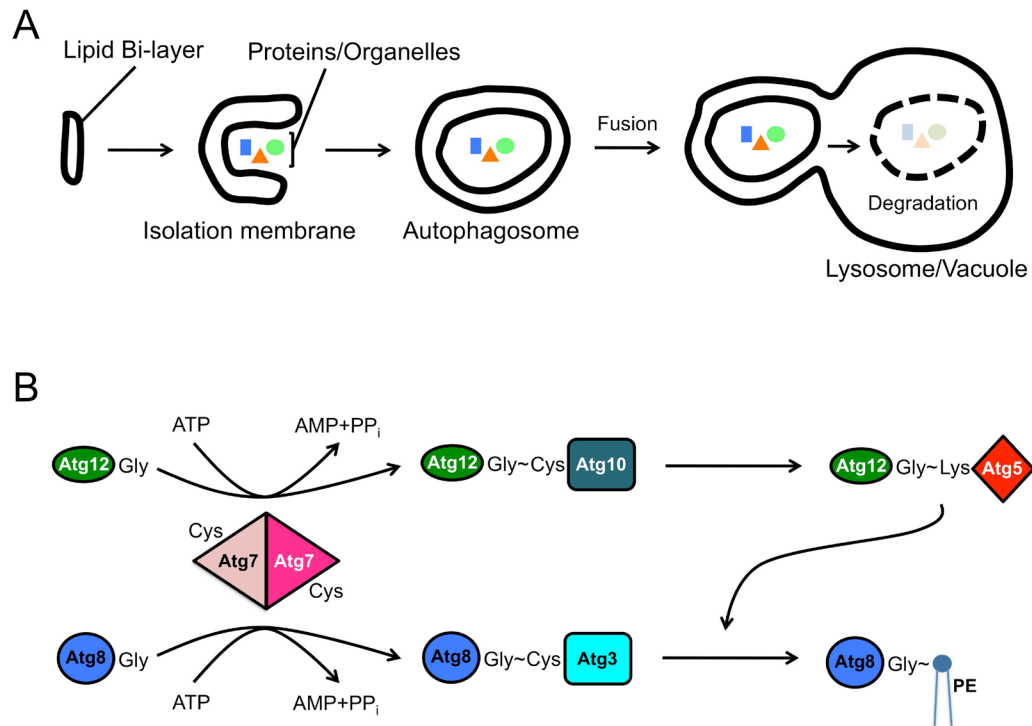


Figure 1-7. Overview of autophagy and its two UBL systems.

(A) Autophagy begins with the formation of a double membrane structure, that expands and starts to engulf portion of the cytoplasm (isolation membrane), This structure then encloses upon itself, forming an autophagosome. The autophagosome then fuses with the lysosome/vacuole where its contents undergo degradation by the resident hydrolases. (B) The Atg12 and Atg8 UBL systems are part of the core autophagic machinery. Atg12 and Atg8 are activated by a common E1, Atg7, and transferred to their E2s, Atg10 and Atg3 respectively. Atg12 then gets conjugated to its final target, Atg5, forming Atg12~Atg5 complex. Atg8 gets transferred to a molecule of phosphatidylethanolamine (PE). This transfer of Atg8, from Atg3 onto PE, is catalyzed by the Atg12~Atg5 conjugate.

Autophagy is mediated by Atg (autophagy-related) gene products. At least 31 Atg genes have been identified to date, of which, 18 are considered core components that are required for autophagosome formation [62-64]. Within this core set of proteins lie two dedicated UBL systems specific for autophagy: the Atg8 and Atg12 system [65-67]. Briefly, Atg8 and Atg12 are activated by a shared E1, Atg7, following which they are thiolester bound to Atg7's catalytic cysteine. Atg8 is then transferred to its specific E2, Atg3 and finally transferred to the head group of phosphatidylethanolamine (PE) – a major component of eukaryotic membranes – to form an Atg8~PE conjugate. Atg12 on the other hand, is transferred to its own specific E2, Atg10 and then to its best-known target, Atg5, forming an Atg12~Atg5 conjugate (Figure 1-7B) [65]. Both Atg8~PE and Atg5~Atg12 are localized to autophagic membranes and required for autophagosome formation. Mutation of any part of the conjugate or protein involved in forming the conjugate negatively affects autophagosome formation [68]. Since the focus in the work described herein is on the Atg8 system, the remainder of this chapter will focus exclusively on Atg8 pathway.

Atg8 – a UBL with a Unique Target and Function

Atg8 is a 13.5 kDa protein that has low sequence homology to ubiquitin (~22% identity) but adopts a ubiquitin fold (Figure 1-2) [69, 70]. Part of the extra molecular weight of Atg8 (ubiquitin has a MW of ~9 kDa) comes from an extra N-terminal helical domain that is not part of the ubiquitin fold. LC3 is the mammalian ortholog of yeast Atg8 [69].

Atg8's final target is PE and studies have shown that the Atg8~PE conjugate plays an important role in autophagosome formation. Specifically, Atg8~PE conjugates have been shown to mediate tethering and hemifusion of PE-congaing liposomes, with increased amounts of the conjugate at the junction of the fusion [71]. No conjugation was seen between a liposome lacking Atg8~PE and one with Atg8~PE [71]. Additionally, the study showed that conjugated Atg8 was able to self-assemble into dimers, trimers and oligomers (while unconjugated Atg8 was monomeric) – a phenomenon that would be required to pull two membranes together for eventual fusion [71]. Interestingly, the N-terminal helical domain of Atg8 plays a role in tethering and hemifusion of membranes [71]. Overall, this information suggests a mechanical role Atg8 in autophagosome formation [37, 64, 71].

Although historically thought to be a non-selective, bulk degradative pathway, autophagy can mediate the delivery of specific cargo to the lysosome. It is well established that Atg8 is recognized by a specific cargo receptor (Atg19) which acts to deliver a specific enzyme, aminopeptidase I (ApeI, a vacuolar hydrolase) to the vacuole in yeast [72, 73]. Similar roles have been proposed for higher eukaryotic Atg8 orthologs and other cargo adaptors in diverse processes such as mitophagy (removal of mitochondria), xenophagy (removal of microorganisms), and the clearance of aggregated proteins from cells [74-76]. This highlights a second role for Atg8 in autophagy and the overall significance of the Atg8 pathway in autophagy.

Atg7 and Atg3 – Atg8's Non-canonical E1 and E2

Atg7 is a common E1 for the two UBL systems present in autophagy and exists as a homodimer [77]. Similar to the canonical E1 enzymes described above, Atg7 possess at its core an adenylation domain. This has been inferred from sequence homology between the C-terminal half of Atg7 compared to MoeB, ThiF and the adenylation domains of canonical E1s [66]. The conservation of a G-X-G-X-X-G nucleotide binding motif and two C-X-X-C Zn²⁺ coordinating motifs are strong indicators of the presence of a typical adenylation domain [64]. Being a homodimer, both adenylation domains are potentially active as is the case with the bacterial MoeB and ThiF homodimeric proteins.

Following the activation of the C-terminus of Atg8, Atg7 next forms a thiolester linkage with Atg8 (Atg7~Atg8) on catalytic cysteine 507 in yeast [67]. Interestingly, unlike canonical E1s that possess a dedicated domain that houses the catalytic cysteine, Atg7's Cys507 likely exists on a loop, similar to MoeB and ThiF, as it lacks additional residues that are necessary to form a cysteine domain.

Subsequent to thiolester formation between the C-terminus of Atg8 and Cys507 of Atg8, Atg7 recruits Atg3, the Atg8-specific E2 [67]. Canonical E1s possess a C-terminal ufd that binds E2s for transfer of UBLs. Atg7 however possess a shorter extension (~60 amino acids) at the C-terminal that appears too short to form the ufd (typically ~100 amino acids), though it does appear to have slight sequence homology to Uba1 (the E1 for ubiquitin) [64]. This raises the question as to how Atg7 recruits Atg3, the E2 for Atg8. One likely candidate for this is Atg7's unique N-terminal domain (~300 amino acids) that lacks sequence homology to any known domain.

Atg3 is a 36 kDa non-canonical E2 that accepts Atg8 from Atg7's catalytic cysteine to form Atg3~Atg8 [64]. Unlike canonical E2s that can have extensions at their N- and/or C-termini, Atg3 has two insertions within the core E2 domain (Figure 1-8), namely a flexible region (Atg3^{FR}) and handle region (Atg3^{HR}) [78]. Both the Atg3^{HR} and Atg3^{FR} are required for Atg8-PE conjugate formation. The Atg3^{HR} is thought to be involved in interacting with Atg8, while Atg3^{FR} interacts with Atg7 [78]. Most of Atg3^{FR} is absent from a crystal structure of yeast full-length Atg3 (PDB code: 2DYT) [78], except a short stretch that adopts a helical conformation. NMR studies show that the FR adopts a random coil conformation in solution [78]. Since Atg7 lacks a complete ufd and Atg3 appears to require the FR region for interaction with Atg7, it will be interesting to decipher how this E1-E2 interaction is mediated.

In the final step of the Atg8 cascade, Atg3 conjugates Atg8 to PE, forming the Atg8~PE conjugate, with a peptide bond is formed between the C-terminus of Atg8 and the amino group of the head moiety of PE [64, 67]. It is here that the two UBL systems (Atg8 and Atg12 systems) display cross-talk with one another. The end product of the Atg12 cascade is Atg12-Atg5 conjugate [66]. This Atg12-Atg5 complex is localized to the expanding autophagosome by another autophagy protein, Atg16 [37, 64, 65]. Here,

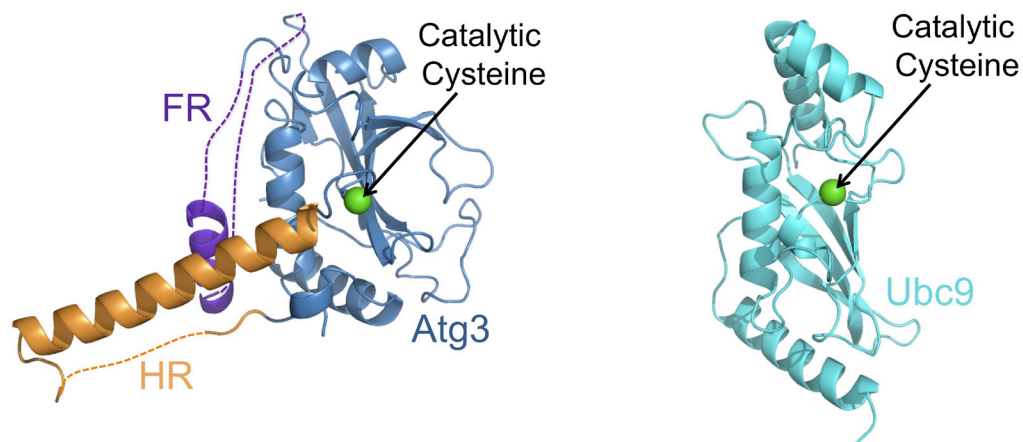


Figure 1-8. Comparison of Atg3 with Ubc9.

Structure of Atg3 (PDB code: 2DYT). The globular core of Atg 3 is shown in blue, with the catalytic cysteine in a green sphere. Atg3 has two insertions, a “flexible region” (FR, purple) and a “handle region” (HR, orange). Dashed lines represent structure with missing electron density. Ubc9 (cyan, PDB code: 1U9A) shown for comparison.

the Atg12-Atg5 complex acts like an “E3”, recruiting the Atg3~Atg8 conjugate to the membrane and speeding up transfer of Atg8 from Atg3 to form Atg8~PE (Figure 1-7B) [64, 79].

The aim of this dissertation study was to gain a deeper understanding into how ubiquitin-like protein E1-E2 interactions are mediated. Therefore, we studied E1-E2 interactions from both, a well-understood canonical pathway and a lesser-studied non-canonical pathway. In Chapter 2, we determined the structure of the Sumo E2, Ubc9, in complex with the ubiquitin-fold domain of Sumo E1 to gain further insights into canonical E1-E2 interactions. In Chapter 3, we performed structural, biophysical and biochemical studies to understand how a non-canonical E1-E2 pair from the autophagy pathway functions. Chapter 4 serves as an overview and a discussion on the future implications of these studies.

CHAPTER 2. CRYSTAL STRUCTURE OF UBA2^{UFD}-UBC9: INSIGHTS INTO E1-E2 INTERACTIONS IN SUMO PATHWAYS*

Introduction

The Small ubiquitin-related modifier (Sumo) protein is a UBL, and like other UBLs, becomes covalently ligated to protein targets (reviewed in [20, 80, 81]). Attachment of Sumo family UBLs are known to alter target functions such as protein-protein interactions, protein-DNA interactions, and subcellular localization (reviewed in [82]). As such, Sumo regulates many important processes, such as signaling, transcription, DNA repair and other stress responses, the cell cycle, and apoptosis [20, 80, 81]. Indeed, the budding yeast Smt3 protein (for simplification referred to as Sumo hereafter) regulates chromosome segregation, formation of the septin ring, and many other aspects of cell division [83-85].

Sumo family members are ligated to proteins via specific E1, E2, and E3 enzymes. The Sumo-specific E1 enzyme (the heterodimeric complex between Uba2 and Aos1 (yeast; termed Sae1-Uba2 in mammals)) initiates the process by first catalyzing adenylation of the Sumo C-terminus, which next becomes linked by a thioester bond to Uba2's catalytic cysteine [86-89]. A transthioesterification reaction ensues during which Sumo is transferred from Uba2 to the catalytic cysteine of the dedicated Sumo E2 conjugating enzyme, Ubc9 [90-93]. Ultimately, either with or without facilitation by a Sumo-pathway-specific E3, Sumo is transferred from the Ubc9 catalytic cysteine to a target lysine ([94] and references therein). In some cases, repeated cycles of Sumo transfer lead to generation of polySumo chains on targets (for review, see [95]). In recent years, structural studies have provided details for many aspects of Sumo E1-E2-E3 conjugation cascades. These include structural understandings of how human Sumo is recognized and activated by Sae1-Uba2 [24, 27], how Ubc9 recognizes targets [36, 94, 96, 97], and how two distinctive E3s function in Sumo ligation [98-100]. Nonetheless, many fundamental aspects of the Sumo cascade remain incompletely understood.

For example, how does the Sumo E1 bind Ubc9? The structure of the human heterodimeric Sae1-Uba2 complex that comprises the Sumo E1 [24] displays conserved E1 domains [23, 101]: a heterodimeric “adenylation” domain comprised of portions of both Sae1 and Uba2, a domain of Uba2 containing the E1 catalytic cysteine, and a C-terminal ubiquitin-fold domain (ufd) from Uba2 that structurally resembles ubiquitin and Sumo [3]. For human Uba2, the catalytic cysteine domain has been shown to make weak interactions with Ubc9 [102]. In addition, for both yeast and human Uba2, the ufd has been implicated in Ubc9 recruitment.

* Chapter 2 adapted with permission. Wang J[#], Taherbhoy AM[#], Hunt HW, Seyedin SN, Miller DW, et al. (2010) Crystal Structure of UBA2^{ufd}-Ubc9: Insights into E1-E2 Interactions in Sumo Pathways. PLoS ONE 5(12): e15805. ([#]*Equal contribution. Contributions: Structure determination by JW, Structural analysis, biochemistry by AMT*)

For canonical UBL pathways, E2s generally appear to bind an E1 ufd [3]. Indeed, crystal structures from the pathway of another UBL, NEDD8, showed previously that the NEDD8 E1's ufd binds NEDD8 E2s, either Ubc12 (also known as UBE2M) or UBE2F, via the E2 N-terminal helix and β 1 β 2-loop [25, 34, 35]. These E1^{ufd}-E2 interactions appear to be conserved across the Sumo, NEDD8, ISG15 and ubiquitin pathways [3]. For example, replacing the UBA6 or UBA7 ufd with that of UBA1 was sufficient to swap E2 specificities, and substituting the E2 UbcH7's N-terminal helix and β 1 β 2-loop regions with that of the E2 UbcH8 swaps E1 specificity [103, 104]. In terms of the Sumo pathway, the human Uba2 ufd was also shown to bind Ubc9 and to play an essential role in human Sae1-Sae2-mediated Sumo transfer to Ubc9 [24, 105]. The ufd of yeast Uba2 is essential for viability [24]. And from the E2 side, budding yeast Ubc9's N-terminal helix and β 1 β 2-loop were also shown to be important for binding to Uba2 and formation of a Ubc9~Sumo complex [28]. Notably, the Ubc9 β 1 β 2-loop is significantly extended relative to other E2s, and has long been recognized as a unique structural feature of Ubc9 [32, 33]. Nonetheless, there is no high-resolution data for E1^{ufd}-E2 interactions other than those for the NEDD8 pathway, despite their importance for the Sumo cascade in particular and UBL transfer in general. Therefore, to obtain a more detailed understanding of Uba2's ufd interactions with Ubc9 and to obtain broad insights into how E1-E2 specificity is established in general, we performed structural analysis of the Uba2^{ufd}-Ubc9 complex from *S. cerevisiae*.

Results and Discussion

Structure of an Isolated Ubiquitin-fold Domain from Uba2 (Uba2^{ufd})

To characterize a domain from yeast Uba2 that binds Ubc9, we determined the crystal structure of the isolated yeast Uba2 ubiquitin-fold domain (ufd), corresponding to residues 439–563, at 1.6 Å resolution (Table 2-1, Figures 2-1 and 2-2). For the most part, residues 440–551 were clearly visible in the electron density with all secondary structures well-defined and for consistency numbered here according to their counterparts in the prior structures of full-length human Uba2 [24]. Only a loop encompassing residues 531–540 displayed weaker electron density precluding building of several side-chains and resulting in high B-factors, presumably due to greater flexibility.

Like ubiquitin and Sumo, the Uba2^{ufd} adopts a modified β -grasp fold. Uba2^{ufd} consists of a twisted 5-stranded antiparallel β -sheet on one side, one 3-turn α -helix on the opposite side of the sheet, and two peripheral short helices. The β -sheet and connecting loops form one V-shaped surface. A second V-shaped surface comes from the edge of β -strand 23, the following loop, and α -helix 37. The two V-like structures are adjacent to each other, and together form a W-shaped surface (Figure 2-3).

Table 2-1. Crystallographic and refinement statistics for Uba2^{ufd} and Uba2^{ufd}-Ubc9.

Accession Codes	Uba2 ^{ufd} (3ONH.pdb)	Uba2 ^{ufd} -Ubc9 (3ONG.pdb)
Data collection		
Beamline	ALS 8.2.1	SERCAT ID
Wavelength (Å)	1.00000	1.00000
Space group	I4	P2 ₁
Cell dimensions		
a, b, c (Å)	a=b=80.198, c=50.733	a=38.465, b=134.038, c=62.012
α , β , γ (°)	$\alpha = \beta = \gamma = 90$	$\alpha = 90$, $\beta = 94.44$, $\gamma = 90$
Resolution (Å)	50-1.6 (1.66-1.6)	45-2.3 (2.38-2.30)
Total reflections	289011	334522
Unique reflections	21034	24002
R _{merge} (%)	9.9 (58.6)	5.6 (26.0)
Average I/ σ	17.8 (3.0)	28.3 (3.5)
Completeness (%)	98.8 (96.4)	85.0 (53.5)
Redundancy	2.9 (2.8)	3.5 (2.9)
Wilson B-factor	15.44	48.9
Refinement		
Resolution range (Å)	40-1.6	38.4-2.3
No. of reflections ($\sigma \geq 0$)	20193	21393
R _{work} (%)	16.0	22.3
R _{free} (%)	17.8	25.1
Number of protein atoms	885	4267
Number of waters	81	38
Average B-factor (protein)	21	72
Average B-factor (water)	30	61
RMSD:		
Bond lengths (Å)	0.006	0.007
Bond angles (°)	0.98	1.08
Ramachandran plot statistics		
Residues in preferred regions (%)	92.8	90.3
Residues in additional allowed regions (%)	6.3	9.7
Residues in disallowed regions (%)	0.9	0

Table 2-1 (continued).

Highest resolution shell is shown in parenthesis. $R_{\text{work}} = \Sigma |F_o - F_c| / \Sigma F_o$. R_{free} is the cross-validation of R-factor, with >5% of the total reflections omitted during model refinement.

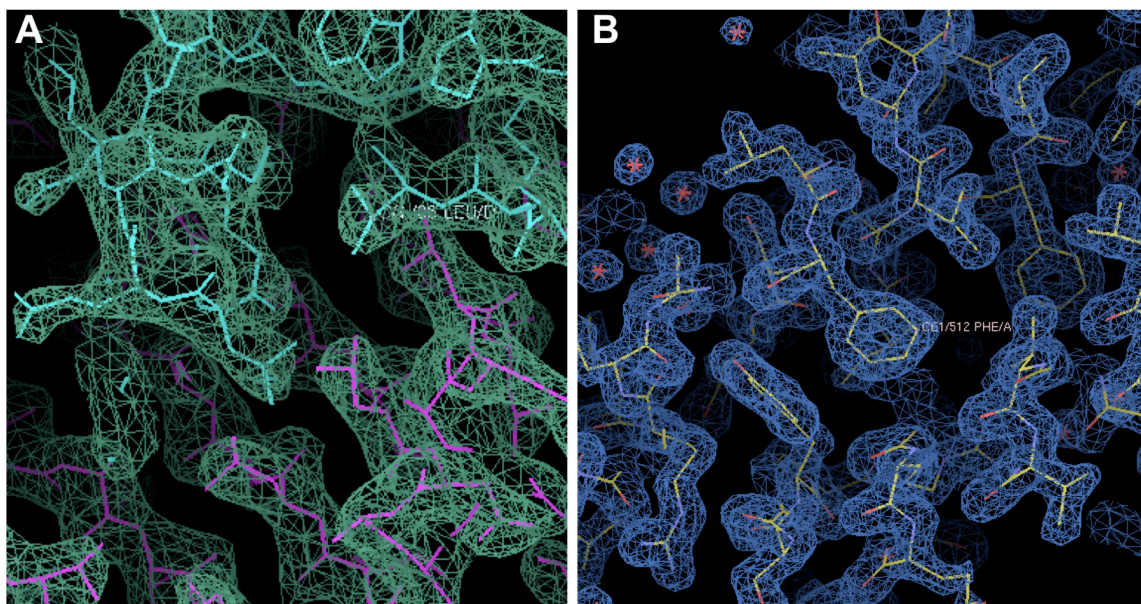
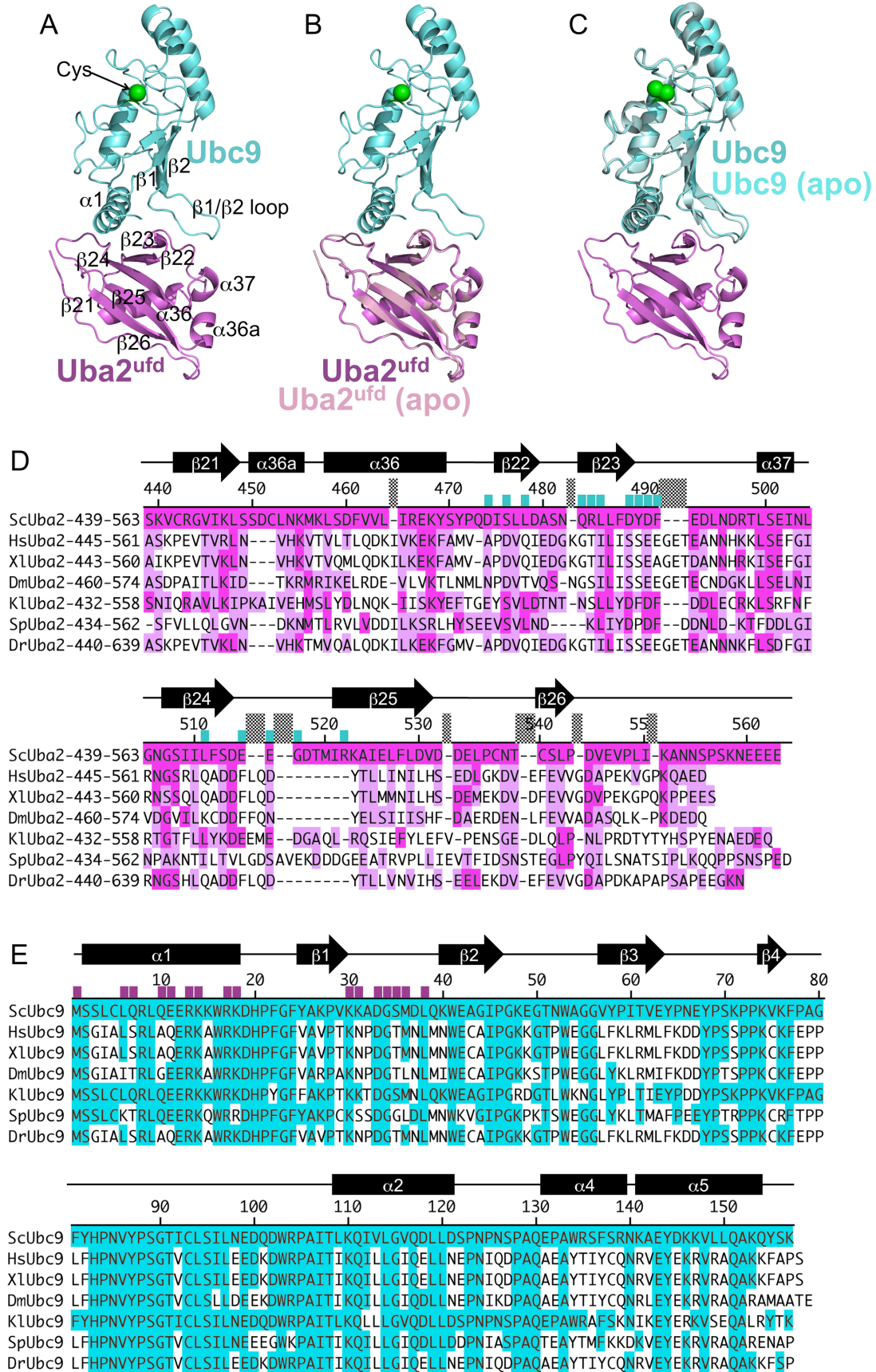


Figure 2-1. Representative electron density.

(A) Final $2F_o - F_c$ electron density contoured at 1.4σ (green mesh) is shown over Uba2^{ufd} (magenta) – Ubc9 (cyan) complex. (B) Final $2F_o - F_c$ electron density contoured at 1.4σ (blue mesh) is shown over Uba2^{ufd} structure.

Figure 2-2. Overall structure of the Uba2^{ufd}-Ubc9 complex.

(A) Cartoon view of the overall structure of the complex with secondary structures numbered as previously described [24, 32, 33]. Uba2^{ufd} is shown in magenta. Ubc9 is shown in cyan. (B) Superposition of free Uba2^{ufd} (pink) superimposed with the Uba2^{ufd}-Ubc9 complex. (C) Superposition of free Ubc9 (seafoam) [106] with the Uba2^{ufd}-Ubc9 complex. (D) Sequence alignment of *Saccharomyces cerevisiae* Uba2^{ufd} sequence (Sc), with the corresponding regions of Uba2 from human (Hs), *Xenopus laevis* (Xl), *Drosophila melanogaster* (Dm), *Kluyveromyces lactis* (Kl), *Schizosaccharomyces pombe* (Sp), and *Danio rerio* (Dr). Alignment was made based on the structures of *S. cerevisiae* Uba2^{ufd} described herein, and the prior structures of human Uba2 [24, 27]. Sequence identity to *S. cerevisiae* Uba2 is highlighted in magenta, and similarity is highlighted in pink, with residues contacting Ubc9 indicated with a cyan bar. Secondary structures are indicated above. (E) Sequence alignment of *Saccharomyces cerevisiae* Ubc9 sequence (Sc), with the corresponding regions of Ubc9 from human (Hs), *Xenopus laevis* (Xl), *Drosophila melanogaster* (Dm), *Kluyveromyces lactis* (Kl), *Schizosaccharomyces pombe* (Sp), and *Danio rerio* (Dr). Sequence identity to *S. cerevisiae* Ubc9 is highlighted in blue, with residues contacting Uba2^{ufd} indicated with a magenta bar. Secondary structures are indicated above.



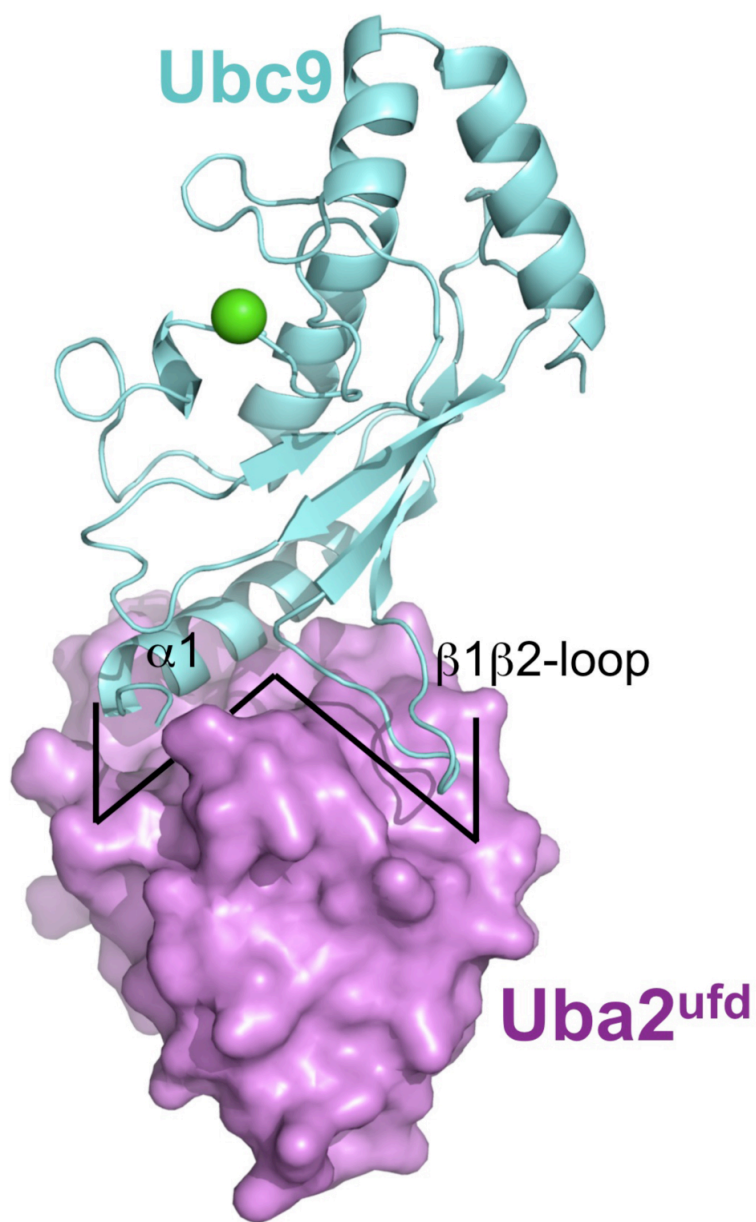


Figure 2-3. Ubc9 binds W-shaped groove in Uba2^{ufd}.

Cartoon view of the overall structure of the complex, with Uba2^{ufd} shown in magenta surface and Ubc9 shown in cyan. The W-shaped groove in Uba2^{ufd} is indicated.

Structure of a Uba2^{ufd}-Ubc9 Complex

In order to understand E1-E2 interactions in a Sumo pathway, we determined the crystal structure of a Uba2^{ufd}-Ubc9 complex from yeast. Phases were obtained by molecular replacement, using the prior structure of yeast Ubc9 and a polyalanine model of human Uba2 residues 449–546 as searchmodels [24, 106]. There are two molecules per asymmetric unit, which superimpose with 0.3 Å root mean square deviation (rmsd) over all alpha carbons, so only one version of the complex is discussed here. The structure of the Uba2^{ufd} is nearly identical in the complex as on its own (0.47 Å rmsd). The structure of Ubc9 superimposes well with the prior structure of free Ubc9 (0.55 Å rmsd), with the most notable difference being slight variation in the position of the β 1 β 2-loop. The Uba2^{ufd}-Ubc9 complex adopts an overall oblong elliptical shape, with the W-shaped surface from Uba2^{ufd} binding the N-terminal α -helix1 and β 1 β 2-loop from Ubc9 (Figures 2-2 and 2-3). This interaction is distal - ~30 Å - to the Ubc9 catalytic cysteine. The Uba2^{ufd}-Ubc9 interface buries 1557 Å² of surface area [107].

Details of Uba2^{ufd}-Ubc9 Interactions

The Uba2^{ufd}-Ubc9 interface is interdigitated, with the two V-shaped grooves from Uba2 binding alternating regions of Ubc9, as in a layered sandwich. The “V” corresponding to the concave Uba2^{ufd} β -sheet holds the Ubc9 N-terminal α -helix1 along its entire length (Figure 2-4A). Ubc9's Leu6 is loosely surrounded by a hydrophobic surface from Uba2's Leu478, Leu485, and Leu511 from three distinct strands. The remainder of the interaction is largely ionic (Figure 2-4A). The Ubc9 N-terminus forms a hydrogen bond with Uba2's Gln483. Ubc9's Gln10 makes a hydrogen bond with Uba2's Ser476 and Glu515. Ubc9's Arg13 makes salt bridges with Uba2's Asp488 and Asp490, and along with Ubc9's Lys14 interacts with Uba2's Tyr489. Ubc9's Lys14 and Arg17 interact with Uba2's Asp474 and Asp490, respectively. At the C-terminus of Ubc9's α -helix1, Lys18 forms a hydrogen bond with the backbone oxygen from Uba2's Gly517.

The Ubc9 β 1 β 2-loop straddles the Uba2 β -strand23. The tip of the Ubc9 β 1 β 2-loop fits in the second “V” shaped groove generated by Uba2 β -strand23, α -helix37, and the intervening loop. Ubc9's Lys30 from one face of the β 1 β 2-loop forms a salt-bridge with Uba2's Asp493. On the other side, Ubc9's Asp33 interacts with Uba2's Arg484. Ubc9's Met36 and Leu38, at the C-terminus of the β 1 β 2-loop, pack against a hydrophobic ridge comprising Uba2's Leu478, Leu485, and Phe491 in the first “V”.

The structure helps explain the deleterious effects of previously reported mutants in yeast Ubc9, which showed that residue substitutions in place of Lys14, or a combination of Lys14, Arg17, and Lys18 from the N-terminal helix, or deleting a portion of the β 1 β 2-loop, hinders binding to full-length Aosl-Uba2 [28]. We further tested the effects of mutating additional key interface residues. Consistent with the structure,

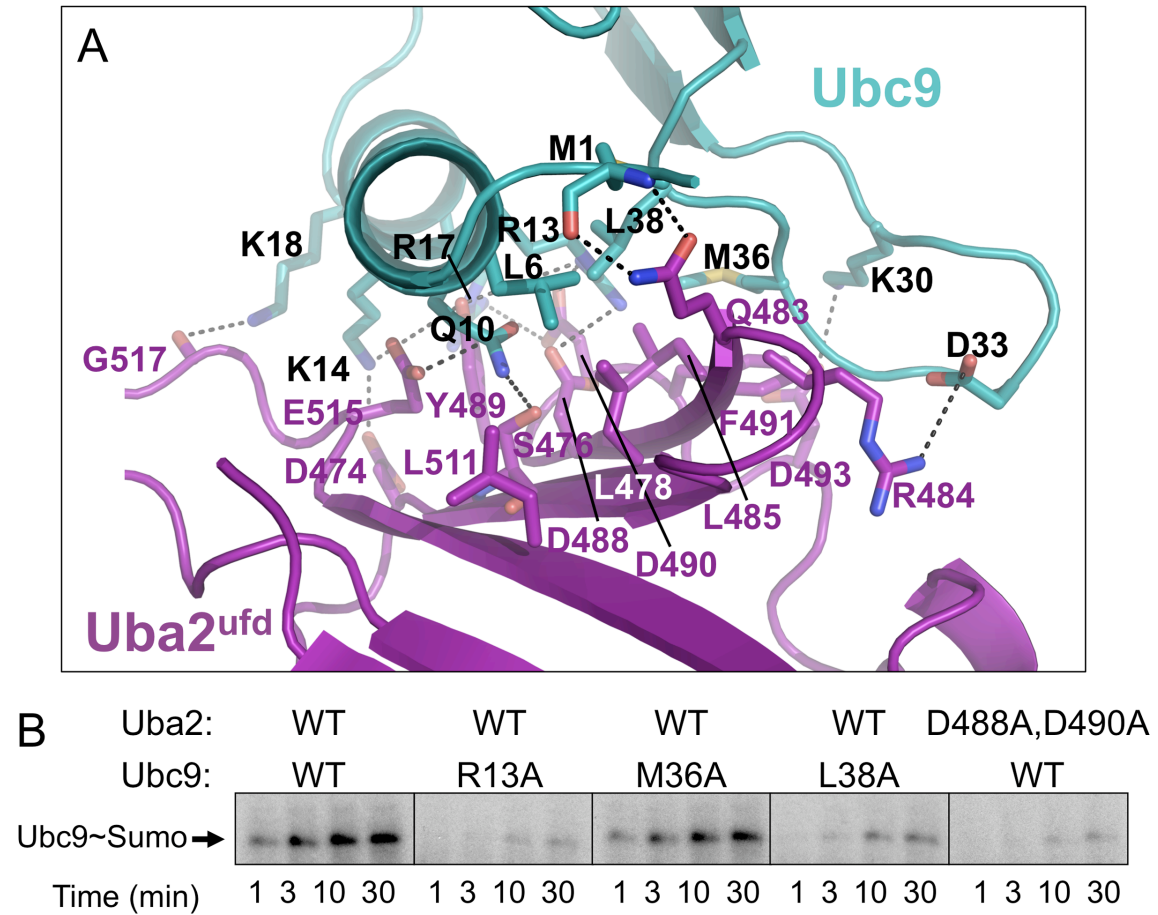


Figure 2-4. Details of the Uba2^{ufd}-Ubc9 interface.

(A) Close-up view of the interface between Uba2^{ufd} and Ubc9. Uba2^{ufd} is shown in magenta. Ubc9 is shown in cyan. Nitrogens are colored blue. Oxygens are colored red. Electrostatic interactions are indicated with dashes. (B) Effects of mutations in Uba2 or Ubc9 at the structurally-observed interface. Shown are autoradiograms monitoring ³²P-labeled yeast Sumo, for time-courses of Aosl-Uba2-catalyzed generation of Ubc9~ [³²P]-Sumo thioester conjugate. The wild-type (WT) or mutant version of Uba2 or Ubc9 used in each panel is indicated above.

individual Ala substitutions in place of Ubc9's Arg13, Met36, or Leu38 hinders Aosl-Uba2-mediated generation of a Ubc9~Sumo complex (Figures 2-4B and 2-5). Similar results are observed for mutating the key acidic patch residues Asp488 and 490 from Uba2 (Figures 2-4B and 2-5).

Conservation of Ubc9 and Uba2 Interacting Residues Across Species

Several structural features and individual amino acids involved in the Uba2^{ufd}-Ubc9 interaction are conserved among Ubc9s and Uba2s from yeast to humans. Yeast and human Ubc9 share 56% sequence identity (Figure 2-2D). Accordingly, the Uba2-binding side-chains from the Ubc9 α -helix1 are 100% conserved as basic residues. Also, the β 1 β 2-loop structure is conserved among Ubc9s, but is uniquely extended relative to E2s for other UBLs. Here, Asp33 is conserved as acidic, and Met36 and Leu38 are conserved as hydrophobics.

The sequence similarity is reflected by extensive structural similarity between yeast and human Ubc9 (0.53 Å C α rmsd, previously noted for other structures in [106]). While the many Ubc9 structures, from different organisms and in different complexes, do generally superimpose well, the primary site of very minor differences is in the orientation of the β 1 β 2-loop. It is possible that limited flexibility of this unique Ubc9 structure may play a role in binding to the Uba2 ufd.

The sequences between Uba2 ufd s are generally less conserved (Figures 2-2E, 2-6A and 2-6B). Despite only 17% sequence identity, a search of the Protein Data Bank using DALI identifies the ufd region from a prior structure of full-length human Uba2 (also called Sae2) as having the highest degree of structural similarity with the yeast Uba2^{ufd} (C α rmsd 1.85 Å) [24, 27, 108]. The main differences are localized to two regions, which are distal to the Ubc9-binding site. The first difference is the short α -helix36a, which corresponds to an insertion in the sequence of yeast Uba2 between β -strand21 and α -helix 36 (Figures 2-2E, 2-6A and 2-6B). The second is at the C-terminal portion of the domain where the sequences are most divergent: the loop preceding β -strand24, the loop between β -strands24 and 25, and the region following β -strand25 that is disordered in human Uba2 but in the yeast Uba2^{ufd} is ordered and forms a more extended structure. Nonetheless, the striking similarity of the overall structures of yeast and human Uba2 ufd s supports the notion that the structurally observed yeast Uba2^{ufd}-Ubc9 interactions will be preserved across species. Indeed, some key interface residues are relatively conserved, including Ser476 (Ser or Asp across species), Arg484 (charged or polar), Leu485 (Leu, Ile, or Val across species), and Asp488 and Asp490 (Ser or Asp, and Asp or Glu across species, respectively) (Figure 2-2E). Furthermore, human Ubc9 was shown to cause chemical shift perturbations in the corresponding region of human Uba2^{ufd} [105].

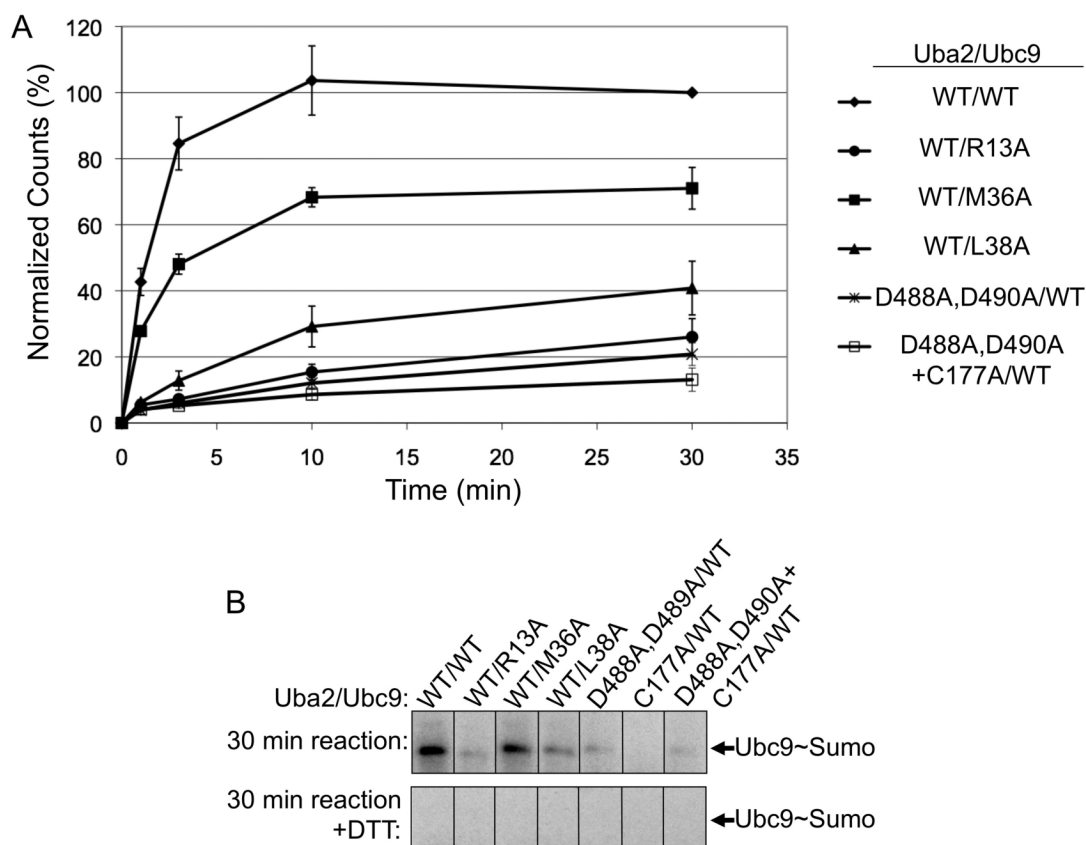


Figure 2-5. Effects of mutations in either yeast Uba2 or Ubc9 on forming the thioester-linked Ubc9~Sumo thioester conjugate.

(A) Amount of Ubc9~ [32 P]Sumo thioester formed over time. Counts were normalized by comparison to the amount of the Ubc9~ [32 P]Sumo reaction product generated for wild-type Uba2 and Ubc9 enzymes at 30 minutes. Error bars represent standard error from experiments performed three independent times. (B) 30-minute time points for reactions shown in A, but treated with DTT prior to SDS-PAGE.

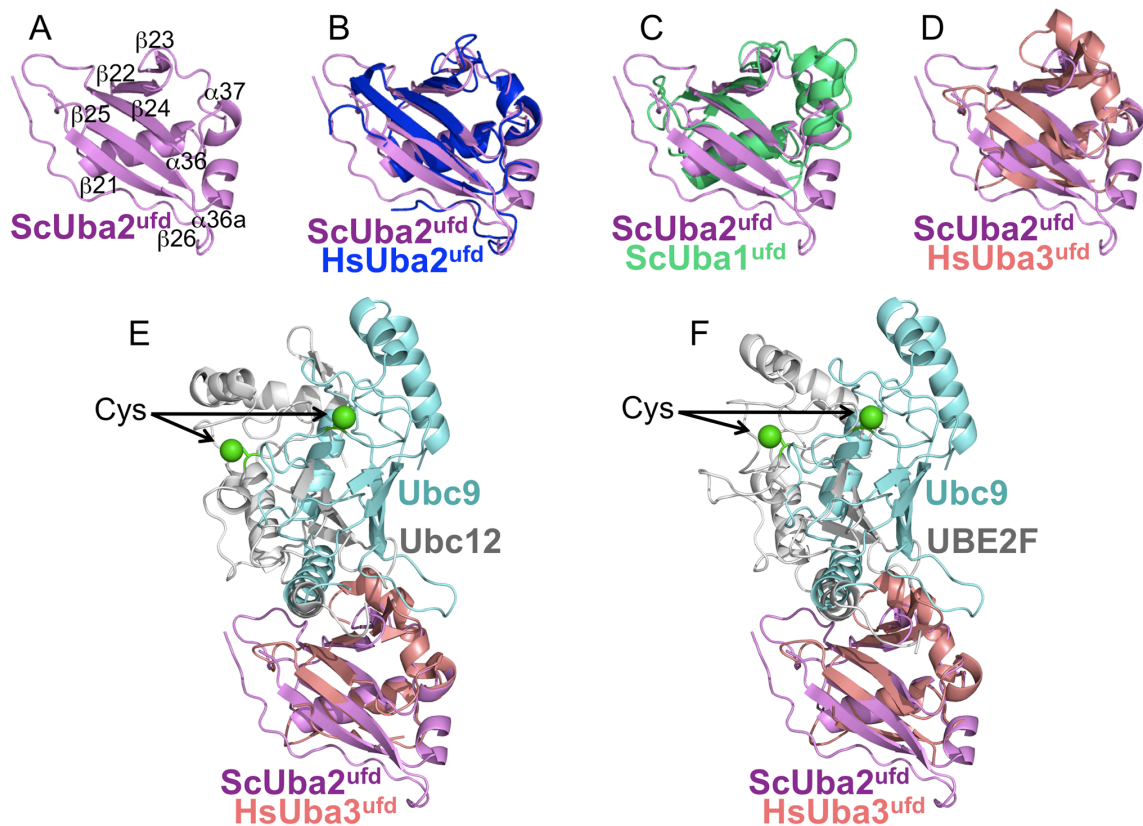


Figure 2-6. Specificity of E1^{ufd}-E2 interactions.

(A) Structure of isolated *S. cerevisiae* Uba2^{ufd}, shown in magenta. (B) Structure of isolated *S. cerevisiae* Uba2^{ufd} superimposed on the corresponding region of human Uba2 (1Y8Q.pdb, blue) [24]. (C) Structure of isolated *S. cerevisiae* Uba2^{ufd} superimposed on the corresponding region of *S. cerevisiae* Uba1, the E1 for ubiquitin (3CMM.pdb, green) [101]. (D) Structure of isolated *S. cerevisiae* Uba2^{ufd} superimposed on the corresponding region of human Uba3 (1Y8X.pdb, peach) [25]. (E) Comparison of E1-E2 interactions in Sumo and NEDD8 pathways by superimposing the E1 ufd from the Uba2^{ufd}-Ubc9 complex structure with a prior human Uba3^{ufd}-Ubc12 (grey) structure (1Y8X.pdb) [25]. (F) Comparison of E1-E2 interactions in Sumo and NEDD8 pathways by superimposing the E1 ufd from the Uba2^{ufd}-Ubc9 complex structure with a prior human Uba3^{ufd}-UBE2F (grey) structure (3FN1.pdb) [35].

Comparison of Uba2^{ufd}-Ubc9 with Prior E1^{ufd} Structures from Other Ubiquitin-like Protein Pathways

To gain insights into similarities and differences among E1 ubiquitin-fold domains, we compared the yeast Uba2^{ufd}-Ubc9 crystal structure with the corresponding regions of crystal structures of yeast Uba1 (E1 for ubiquitin) and human Uba3 (E1 for NEDD8) (Figures 2-6A, 2-6C, 2-6D and 2-7). The yeast Uba2^{ufd} displays overall structural similarity to the ufd's from Uba1 and Uba3 (2.4 and 1.6 Å rmsd, respectively), including the W-shaped surface. Furthermore, structural comparison of Uba2^{ufd}-Ubc9 with prior structures of Uba3 in complex with NEDD8 E2s (Ubc12 or UBE2F) reveal that the Sumo and NEDD8 pathways use parallel E1^{ufd} and E2 surfaces to mediate interactions (Figures 2-6E and 2-6F).

Despite their overall common mode of binding, an intriguing difference is that the relative positions of the E1^{ufd} and E2 catalytic cysteines are considerably offset between Uba2^{ufd}-Ubc9 and Uba3^{ufd}-Ubc12 or Uba3^{ufd}-UBE2F (Figures 2-6E and 2-6F). The detailed interactions between Uba2 and Ubc9 described above (Figure 2-4A) also differ substantially from corresponding contacts in the NEDD8 pathway [25, 35]. Furthermore, global structural differences between Uba2 and other E1 ufd's, and between Ubc9 and other E2s help explain why Ubc9 does not bind E1s for other UBLs and vice-versa. In the Uba1 ufd the region corresponding to the Uba2^{ufd}'s short α -helix37 is extended by a complete turn, and the preceding loop in Uba2^{ufd} corresponds to a 2-turn helix in Uba1 (Figure 2-6C) [101]. This latter helix is even longer in Uba3's ufd (Figure 2-6D) [23, 25]. These Uba1 and UBA3 helical insertions would clash with the Uba2-binding basic residues from Ubc9's α -helix1 and extended β 1 β 2 loop. Thus, several structural differences between E1 ufd's, and corresponding distinctions between their cognate E2s, may help establish specificity of E1-E2 interactions.

Docking Uba2^{ufd}-Ubc9 onto Prior Full-length Human Aos1-Uba2 Structures

In order to gain insights into Sumo transfer from Uba2 to Ubc9, we superimposed the Uba2^{ufd}-Ubc9 structure onto the corresponding region of human Uba2 from several recent crystal structures [24, 27]. The overall locations of domains, including the ufd, are generally similar in structures of human Aos1-Uba2-MgATP (also called Sae1-Sae2-MgATP, 1Y8Q.pdb), Aos1-Uba2-Sumo-MgATP (1Y8R.pdb), and Aos1-Uba2-Sumo~AMSN (a chemical modification mimicking Sumo C-terminal adenylation, 3KYC.pdb) [24, 27]. The ufd is also in a similar relative orientation in a recent structure of human Aos1-Uba2~Sumo~AVSN, which mimics the tetrahedral intermediate during Uba2's catalytic cysteine's attack on the C-terminally adenylated Sumo (3KYD.pdb), so only the comparison to Aos1-Uba2-Sumo~AMSN is shown in Figure 2-8A. All the structures of full-length Uba2 lack Ubc9, and represent states of Aos1-Uba2 prior to Sumo transfer to Ubc9 [24]. All docking models on full-length Uba2 would place Ubc9's catalytic cysteine distal to and facing the opposite direction of Uba2's catalytic cysteine (Figure 2-8A).



Figure 2-7. Sequence comparisons of Ubc9, Uba2^{ufd}, and corresponding regions of E2s and E1s for other UBLs.

(A) Sequence alignment of *Saccharomyces cerevisiae* Ubc9 sequence (Sc), with Ubc9 from human (Hs), and the catalytic core domain regions of the human E2s for NEDD8 (Ubc12 and UBE2F). Sequences were aligned based on structures. Secondary structures are indicated above. (B) Sequence alignment of *Saccharomyces cerevisiae* Uba2^{ufd} sequence (Sc), with the corresponding regions of Uba2 from human (Hs), and the ufd domains from E1s for ubiquitin (Uba1) and NEDD8 (Uba3). Sequences were aligned based on structures. Secondary structures from the Sc Uba2^{ufd} structure are indicated above.

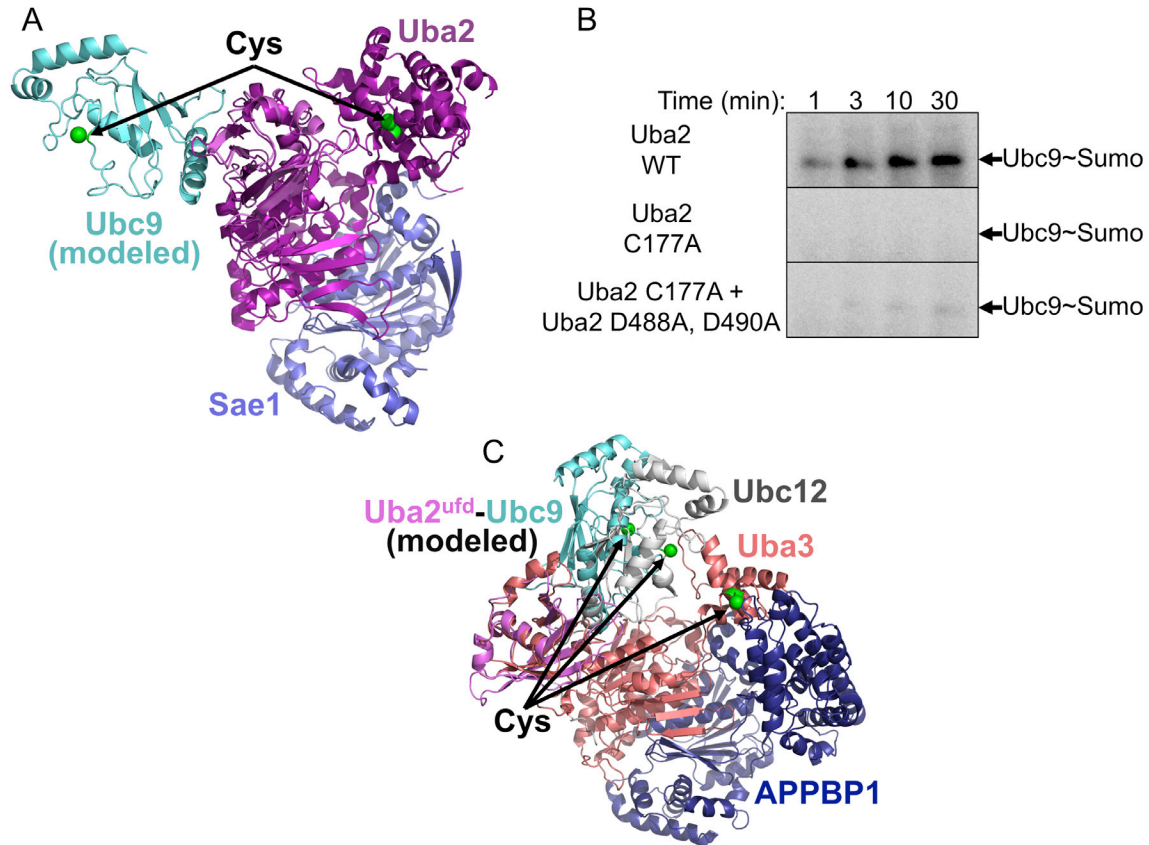


Figure 2-8. Implications of Uba2^{ufd}-Ubc9 structure for Smt3 transfer from Uba2 to Ubc9.

(A) Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan), with Uba2^{ufd} superimposed on the corresponding region of human Uba2 from the Sae1-Uba2-Sumo1~AMSN complex (3KYC.pdb) [27]. Sae1 is shown in slate, human Uba2 is shown in purple, and Sumo~AMSN is not shown for simplification. Catalytic cysteine residues of Ubc9 and Uba2 are shown in green. (B) Autoradiograms monitoring ³²P-labeled Sumo, for time-courses of Aos1-Uba2-catalyzed generation of Ubc9~ [³²P]-Sumo thioester conjugate. WT – wild-type; Uba2 C177A – cysteine-to-alanine mutation at Uba2 catalytic Cys177. Uba2 D488A,D490A – aspartate-to-alanine mutation at Uba2's Ubc9-interacting Asp488 and Asp490. (C) Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan), with Uba2^{ufd} superimposed on the corresponding region of human Uba3 (peach) from the APPBP1 (blue) -Uba3~NEDD8-NEDD8-Ubc12(C/A)-MgATP complex [32]. For simplification, the NEDD8 molecules are not shown, and the E2 active site positions (an Ala mutation in Ubc12) are labeled with green spheres, and “Cys” to denote their locations.

To address the possibility of whether Sumo thioester-linked to the catalytic Cys177 of one Uba2 can be transferred to Ubc9 bound to another Uba2, we examined generation of the Ubc9~Sumo thioester-linked conjugate upon mixing two defective mutant versions of the Sumo E1. In one mutant, the Uba2 catalytic Cys177 was mutated to Ala to prevent formation of a Uba2~Sumo intermediate, but the *ufd* was wild-type. The other retained ability to form a thioester linkage with Sumo at Cys177, but harbored Ala mutations in the Ubc9-binding *ufd* residues Asp488 and Asp490. On their own, each mutant is defective for generating the thioester-linked Ubc9~Sumo conjugate. Mixing the two E1s does not rescue the other's defect (Figure 2-8B).

Thus, it seems likely that Uba2 undergoes a conformational change to bring the Uba2 and Ubc9 catalytic cysteines together for Sumo transfer. Previous structural studies of the NEDD8 E1, which showed dramatic *ufd* reorientation, may provide some insights into such a conformational change. Superimposing Uba2^{*ufd*}-Ubc9 on the Uba3^{*ufd*} in a structure of the NEDD8 E1 complexed with two molecules of NEDD8 and a catalytically-inactive E2 (Ubc12), which represents an intermediate en route to UBL transfer from E1 to E2, shows how *ufd* rotation might allow Ubc9 to be reoriented toward Uba2's catalytic cysteine [34]. However, even upon rotation of the Uba2 *ufd* to the position observed in the NEDD8 E1, a significant E1-to-E2 cysteine-to-cysteine gap remains (Figure 2-8C). It is possible that Uba2's *ufd* undergoes an even greater rotation during Sumo transfer. Another excellent possibility for how the gap between the Uba2 and Ubc9 catalytic cysteines could be closed comes from the recent landmark finding that the Uba2 catalytic cysteine domain can undergo dramatic reorientation and remodeling [27]. It seems likely that the proposed combination of *ufd* and catalytic cysteine domain rotations bring the Uba2 and Ubc9 cysteines together.

Implications for E1-E2-E3 Cascades in the Sumo Pathway

E2s play the central role in UBL cascades. Thus, after receiving Sumo from Uba2, Ubc9 ultimately transfers Sumo to a target. This is often facilitated by an E3. Two types of Sumo E3s have been structurally characterized: (1) a neither-HECT-nor-RING E3, as in RanBP2/Nup358, which binds both Ubc9 and its covalently-linked Sumo to optimally orient the thioester bond for Sumo transfer to a substrate Lys [98, 109, 110] and (2) Siz/PIAS E3s, which contain a SP-RING domain that structurally resembles RING E3s utilized by the ubiquitin and NEDD8 pathway [99, 111]. To gain insights into how Sumo E1-E2-E3-target cascades are organized, we superimposed Ubc9 from our Uba2^{*ufd*}-Ubc9 structure with that in the prior crystal structure of Sumo~RanGAP1-Ubc9-RanBP2/Nup358 [98], and with a structural model of Ubc9-Siz1 based on other E2-RING domain structures [99, 112]. Uba2^{*ufd*} and RanBP2/Nup358 bind overlapping surfaces on Ubc9 [24] (Figure 2-9A). Structural modeling also indicates that Uba2^{*ufd*} would clash with the Siz1 SP-CTD domain, which is C-terminal of the SP-RING domain and is conserved among Siz/PIAS E3s (Figure 2-9B) [99]. Thus, Uba2-mediated formation of a Ubc9~Sumo thioester conjugate may be mutually exclusive with E3-mediated Sumo ligation to targets.

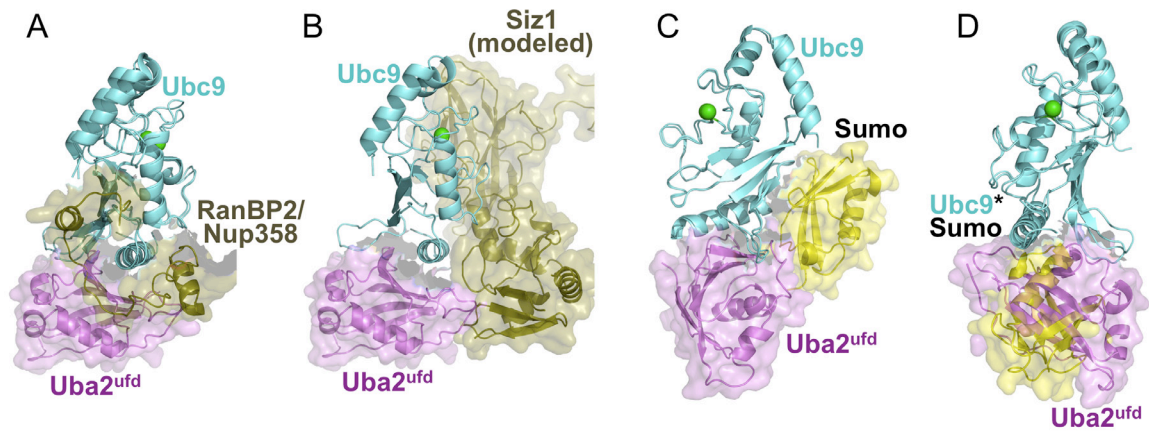


Figure 2-9. Implications of Uba2^{ufd}-Ubc9 structure for Sumo cascades.

(A) Comparison of E1-E2 and E2-E3 interactions in Sumo cascades. Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan), with Ubc9 superimposed on the previous structure of human Ubc9 (cyan) bound to RanBP2/Nup358 (olive) from the complex with RanGAP-1 (not shown) and Sumo-1 (not shown) [98]. Uba2^{ufd} and a portion of RanBP2/Nup358 are shown with a semi-transparent surface to highlight the overlapping regions. (B) Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan), with Ubc9 superimposed on a model for *S. cerevisiae* Ubc9-Siz1 (olive), based on previous structures of Siz1 and c-Cbl-UbcH7 [99, 112]. Briefly, the RING domain c-Cbl and UbcH7 were docked onto the SP-RING domain of Siz1, and then Ubc9 was modeled in place of UbcH7. Uba2^{ufd} and Siz1 are shown with a semi-transparent surface to highlight the overlapping regions. (C) Comparison of E1-E2 and noncovalent E2-UBL interactions in Sumo cascades. Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan) superimposed in previous structure of noncovalent Ubc9-Sumo (cyan, yellow) complex from *S. cerevisiae* [113]. Uba2^{ufd} and Sumo are shown with a semi-transparent surface to highlight the overlapping regions. (D) Structure of *S. cerevisiae* Uba2^{ufd}-Ubc9 (magenta, cyan), with Ubc9 superimposed on human Ubc9*Sumo (cyan*yellow) [114]. Uba2^{ufd} and Sumo are shown with a semi-transparent surface to highlight the overlapping regions.

Sumo also binds noncovalently to Ubc9, via the “backside” distal from the E2 catalytic cysteine and this interaction is conserved from yeast to humans [113, 115, 116]. Although the function of this noncovalent Ubc9-Sumo interaction is not well understood, mutant versions of Ubc9 where this interaction is disrupted cannot support yeast viability [113], and are impaired at generating long poly-Sumo chains [115, 116]. Previous biochemical data showed that the noncovalent Sumo interaction competes with Ubc9's interaction with Uba2 [28, 113]. Indeed, superposition of Ubc9 from our complex with Uba2^{ufd} with prior structures of complexes with Sumo reveal mutually exclusive binding to Ubc9's α -helix1, particularly Arg13 and Arg17 (Figure 2-9C). Thus, the Uba2^{ufd} recognizes multipurpose binding sites on Ubc9, which may serve to order association with many different partners in the Sumo cascade.

In a related vein, human Ubc9 is distinct from yeast Ubc9 in becoming modified on Lys14 via an isopeptide bond to human Sumo, forming a covalent complex referred to as Ubc9*Sumo [114]. Ubc9*Sumo can bind E1 and form a thioester-linked complex to another Sumo at its active site, which subsequently can be transferred to targets. Sumoylation of Ubc9 shifts its E3-independent target specificity [114]. For the substrate Sp100, it seems Sp100's Sumo-interacting motif (SIM) is recruited to the Sumo linked to Ubc9's Lys14. Comparison of the crystal structure of human Ubc9*Sumo with that of yeast Uba2^{ufd}-Ubc9 reveals that Uba2 overlaps the position of Sumo in the covalent complex (Figure 2-9D). How then might human Ubc9*Sumo bind Uba2 to become charged with Sumo? Given the high structural similarity between human and yeast Ubc9 [106] and Uba2^{ufd} (Figure 2-6), and prior NMR and mutational data consistent with conservation of our structurally-observed Uba2^{ufd}-Ubc9 interactions in their human counterparts [24, 29, 105], it seems likely that human Uba2 and Ubc9 interact as in our structure. A plausible model would be that there is some flexibility between Ubc9 and Sumo in Ubc9*Sumo such that Sumo's β -grasp domain moves out of the way when Uba2 binds. This is consistent with ~4-fold more human Aos1-Uba2 being required to fully bind Ubc9*Sumo in comparison to Ubc9 in a qualitative interaction study [114]. Additional interactions might stabilize Uba2-Ubc9 complex formation during the transthiolation reaction. Indeed, although quantitative enzyme kinetic comparison was not performed, under some conditions Ubc9 and Ubc9*Sumo show equal formation of thioester-linked intermediates with Sumo [114]. Future studies will be required to understand the dynamics of interactions between Sumo, Ubc9, Uba2, and E3s in transfer cascades.

Conclusions

The structure of the yeast Uba2^{ufd} in complex with Ubc9 provides the first crystallographic insights into E1-E2 contacts in a Sumo pathway. The structure is consistent with mutational data presented previously and herein, which validate these interactions for Sumo conjugation to Ubc9 [24, 28, 29, 102, 105, 106].

The complex structure also provides only the second detailed view of E1^{ufd}-E2 interactions for any UBL, and demonstrates a common overall mode of E2 recruitment by

E1 ufd's: a ufd's "W" (i.e., double-"V") -shaped surface engages an E2's N-terminal helix and $\beta 1\beta 2$ -loop. Specificity within the Sumo pathway is established by many unique favorable interactions between Uba2 and Ubc9. Furthermore, Ubc9 is distinct among E2s as having an extended rigid $\beta 1\beta 2$ -loop. This is recognized by a conserved C-terminal "V" from Uba2^{ufd}, comprised of Uba2 elements spanning from β -strand23 through α -helix37. The Uba2^{ufd} C-terminal "V" is minimized relative to the corresponding regions in ubiquitin and NEDD8 E1s [23, 101], which would be unable to accommodate the Ubc9 $\beta 1\beta 2$ -loop. Structural comparison to Ubc9-E3 structures/models, and to Ubc9-Sumo, demonstrates how E1-E2 interactions are mutually exclusive with forming E2-E3 or noncovalent E2-UBL complexes in Sumo cascades.

Finally, docking the Uba2^{ufd}-Ubc9 structure on prior structures of full-length Uba2 indicate that the Uba2 ufd would need to undergo significant rotation so that the Uba2 and Ubc9 cysteines could face each other. This is also a common feature of previous structures from the NEDD8 cascade [23, 25] [26, 34]. Furthermore, based on conformational changes observed for Uba2 at earlier steps in the Sumo activation cycle, it seems that the Uba2 cysteine domain may also undergo conformational changes to meet the Ubc9 cysteine for Sumo transfer [27]. Along these lines, it is noteworthy that unique features to Uba2s and Ubc9s also establish a distinct orientation of the E2 catalytic cysteine relative to the E1's E2-binding ufd (Figures 2-6E and 2-6F). It will be interesting in the future to understand how UBL-specific E1 conformational changes accommodate such distinct geometries during E1-to-E2 UBL transfer.

Materials and Methods

Cloning, Protein Expression and Purification

Expression constructs were made by standard PCR/ligation procedures, with sequences verified by automated sequencing procedures. Yeast Sumo, Smt3, is referred to as Sumo in the text. For biochemical studies, Smt3/Sumo residues 13–97 were expressed from pGEX-2TK (GE), harboring an N-terminal site for phosphorylation by Protein Kinase A (PKA), and was purified as previously described for other UBLs [35].

Uba2^{ufd} (residues 439–563) was expressed from pGEX-4T1 (GE). BL21(DE3) harboring the expression construct were grown in LB with ampicillin (100 μ g/mL) at 37°C until reaching an OD₆₀₀ of ~1.0. Expression was induced by the addition of IPTG to a final concentration of 0.6 mM, followed by overnight culturing at room temperature. Cells were resuspended in 50 mM Tris-HCl, 0.2 M NaCl, 5 mM DTT, pH 8.0 supplemented with 2.5 mM with PMSF, and lysed by sonication on ice. GST-Uba2^{ufd} was purified by glutathione-affinity chromatography, treated with a 1:100 ratio of thrombin during overnight dialysis into 25 mM Tris-HCl, 0.2 M NaCl, 5 mM DTT, pH 8.0 at 4°C overnight. GST and any remaining GST-Uba2^{ufd} fusion protein were removed by glutathione-affinity. For co-crystallization, Uba2^{ufd} was purified by gel filtration chromatography using a SD75 column (GE) in 25 mM Tris-HCl, 0.15 M NaCl, 5 mM

DTT, pH 7.6, concentrated to 22 mg/ml (Bio-Rad Protein Assay), aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C until further use. For crystallization on its own, Uba2^{ufd} prepared similarly, except that gel filtration was performed in 50 mM Tris-HCl, 0.2 M NaCl, 5 mM DTT, pH 7.6, and Uba2^{ufd} was concentrated to 40 mg/ml.

Ubc9 was expressed from pGEX4T3 (GE) [28] in BL21(DE3) Gold cells. After growth at 37°C until the OD₆₆₀ reached ~ 0.8 , expression was induced by the addition of 1 mL of 0.6 M IPTG/liter of culture, and induction was carried out overnight at 16°C . GST-Ubc9 was purified by glutathione affinity. For biochemical studies, GST-Ubc9 bound to glutathione sepharose was treated with thrombin overnight at 4°C to release wild-type and mutant versions of Ubc9. Wild-type and mutant versions of Ubc9 were concentrated in a final buffer of 50 mM Tris-HCl pH 7.6, 200 mM NaCl, 1 mM DTT, aliquotted, flash-frozen, and stored until use. For crystallography, GST-Ubc9 was treated with a 1:30 ratio of thrombin overnight at 16°C . Ubc9 was purified to homogeneity by cation exchange with a homemade column packed with ResourceS resin (GE) using a NaCl gradient in 20 mM HEPES, 1 mM DTT, pH 7.0, concentrated to 27 mg/ml, aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C until further use.

For biochemical assays, the expression construct for Sumo E1 (the heterodimeric Aos1-Uba2 complex) was described previously [28]. BL21(DE3) Gold cells harboring the expression construct were grown in LB with ampicillin (100 $\mu\text{g}/\text{mL}$) at 37°C until reaching an OD₆₀₀ of ~ 0.8 . After cooling to 16°C for 45 minutes, expression was induced by the addition of 1 mL of 0.2 M IPTG/liter of culture, shaking overnight at 16°C . GST-Aos1-Uba2 was purified by glutathione affinity chromatography, and eluted protein was cleaved overnight with thrombin at 4°C . Aos1-Uba2 was further purified by anion exchange chromatography, concentrated in a final buffer of 50 mM Tris-HCl pH 7.6, ~ 290 mM NaCl, 1 mM DTT, aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C until further use.

Crystallization, Data Collection, and Structure Determination and Refinement

Initial crystallization screening was performed with a Mosquito crystallization robot (TTP Labtech) with commercial 96-well screens. After testing several conditions (chemical and temperature) and Uba2^{ufd}-Ubc9 complexes (from human and yeast with different ufd domain boundaries), crystals were obtained at 4°C in the Index HTS condition G2 (Hampton Research), for a 1.1:1 volume:volume mixture of Uba2^{ufd} at 22 mg/ml and Ubc9 at 27 mg/ml. A diffraction-quality crystal was obtained after multiple cycles of streak-seeding into 20% (w/v) PEG3350, 0.1 M Bis-Tris pH 5.5, 0.2 M Li₂SO₄, and harvested and flash-frozen in mother liquor supplemented with 20% (v/v) glycerol. The crystals belong to space-group P2₁. Data to 2.3 Å resolution were obtained at the SERCAT-ID beamline, and processed with HKL2000 [117]. Initial models for the complex were obtained by performing molecular replacement in the P1 space group using PHASER [118], after searching for 4 copies of the structure of yeast Ubc9 (2GJD.pdb, Chain A) [106] and of a polyAla model of human Uba2 residues 449–546 from 1Y8Q.pdb, Chain C [24]. One of the 4 complexes from this solution was subsequently

used as a search model to find two Uba2^{ufd}-Ubc9 complexes per asymmetric unit in P2₁. The model was built using Coot [119] and refined using TLS parameters in REFMAC [120]. Due to low sequence similarity with the human Uba2^{ufd}, we wished to obtain additional data to confirm the structure in regions where side-chains were poorly visible. We thus obtained crystals for the isolated Uba2^{ufd} at 4°C in the Ammonium Sulfate HTS condition H8 (Qiagen). After optimization in manual trays, a crystal was obtained in 3 M (NH₄)₂SO₄, 1% (v/v) MPD, and harvested and flash-frozen in 3 M AmSO₄, 1% (v/v) MPD, 25% (v/v) glycerol. Data to 1.6 Å resolution were obtained at the SERCAT-ID beamline, and processed with HKL2000 [117]. Examination of the data with Phenix.xtriage indicated the space group to be I₄, with nearly perfect merohedral twinning and a twin operator of -k,-h,-l. The structure was obtained by molecular replacement using PHASER [118], with the Uba2^{ufd} from the partially-refined Uba2^{ufd}-Ubc9 complex structure as a search model for one molecule in the asymmetric unit. After rebuilding using Coot [119] and refinement including twin laws using Phenix [121], the isolated Uba2^{ufd} structure was then used as a guide for final refinement of the Uba2^{ufd}-Ubc9 structure. Data collection and refinement statistics are provided in Table 2-1. The single Ramachandran outlier in the high resolution Uba2^{ufd} structure corresponds to Gly517, with the conformation supported by the electron density.

Assays for Formation of a Ubc9~Sumo Thioester Conjugate

Sumo was labeled with ³²P at an N-terminal PKA site using PKA (New England Biolabs) and [γ-³²P]ATP as described previously for other UBLs [35]. Reactions were performed at 18°C, which is the ambient room temperature for our dedicated room for use of radioactivity, in 10 µl reaction volumes containing 50 mM Tris-HCl (pH 7.6), 5 mM ATP, 10 mM MgCl₂, 2 mg/ml BSA, 5 mM creatine phosphate, 0.3 U/ml creatine kinase, 0.3 U/ml inorganic pyrophosphatase, 5 nM Aosl-Uba2, 250 nM Ubc9 or the indicated mutant, and 10 µM [γ-³²P]-Sumo. Reactions were quenched at 1, 3, 10 and 30 minutes by the addition of an equal volume of 2X SDS sample buffer, resolved on 15% SDS-PAGE gels, dried, and exposed to film. In the assay containing both Aosl-Uba2 (Cys177Ala) and Aosl-Uba2 (Asp488Ala and Asp490Ala), 5 nM of both mutant E1s was used.

CHAPTER 3. STRUCTURE OF AN ATG7-ATG3 SUB-COMPLEX REVEALS A DISTINCTIVE INTERACTION AND MECHANISM FOR TRANSTHIOLATION UNIQUE TO THE AUTOPHAGY PATHWAY

Introduction

In response to a variety of pathogenic states or regulatory cues such as low cellular energy, eukaryotic cells initiate the degradative process of macroautophagy (hereafter autophagy, reviewed in [37-39, 122]). During autophagy, a double-membrane vesicle (termed the autophagosome) engulfs a portion of the cytoplasm and fuses with a lysosome or vacuole. Ultimately, the cytosolic contents inside the autophagosome are degraded and recycled (reviewed in [37, 123]). At least 18 distinct so-called “Atg” proteins are key core components for autophagic membrane formation [37, 64, 124]. These include the ubiquitin-like protein (UBL) Atg8 in yeast, which has six orthologs in mammals including LC3 [3, 67]. Atg8 is unique among all UBLs in being ligated to a lipid, phosphatidylethanolamine (PE) [67, 125]. Atg8~PE (~ = covalent bond) adducts are incorporated into the growing autophagosomal membrane, critical for its expansion [71].

As with other UBLs, Atg8 becomes ligated to its target via a multistep process (reviewed in [65]). First, the E1 enzyme Atg7 catalyzes adenylation of Atg8’s C-terminus. Next, the resulting Atg8-acyl adenylate is attacked by Atg7’s catalytic cysteine (Cys) to generate an Atg7~Atg8 intermediate, which is linked by a thioester bond. A transthioylation occurs whereby Atg8’s C-terminus is transferred from the Atg7 catalytic Cys to that of an E2 enzyme, Atg3, generating a thioester-linked Atg3~Atg8 conjugate. Finally, Atg8 is transferred from the Atg3 catalytic Cys to PE [67]. Although Atg7 and Atg3 are sufficient to mediate Atg8 ligation *in vitro*, this is stimulated by an unusual multiprotein E3 containing the Atg12~Atg5 complex [79]. This is partly composed of Atg12, another UBL involved in autophagy. Like Atg8, Atg12 is activated by Atg7, although it has its own E2 (Atg10) and target (Atg5) [66].

Models for the initial steps of Atg8 activation may be extrapolated from (1) structural similarity of Atg8 to canonical UBLs such as ubiquitin, NEDD8, and Sumo, and bacterial antecedents MoaD and ThiS (Figure 3-1) [64, 69], and (2) sequence homology between roughly half of Atg7 to the adenylation domain found in canonical E1s and homodimeric bacterial E1 antecedents such as MoeB and ThiF (Figure 3-2) [66] [64]. On this basis, it is possible to generate a homology model of Atg8 bound to the adenylation active site of Atg7 (Figures 3-1B and 3-3). However, detailed mechanisms of Atg8 conjugation for the steps following the initial adenylation reaction cannot be extrapolated from knowledge of canonical UBL pathways due to numerous fundamental differences. First, like *E. coli* ThiF, Atg7’s catalytic Cys resides in a short loop, which is near the adenylation active site. By contrast, in canonical E1s such as for ubiquitin, Nedd8, and Sumo, the catalytic Cys is in a distinct domain, which undergoes a conformational change to move close to the adenylation active site for becoming covalently linked to the C-terminus of a canonical UBL [3, 27]. Another difference is that Atg7 is a homodimer [77]. Although roles of homodimerization are unknown, there are

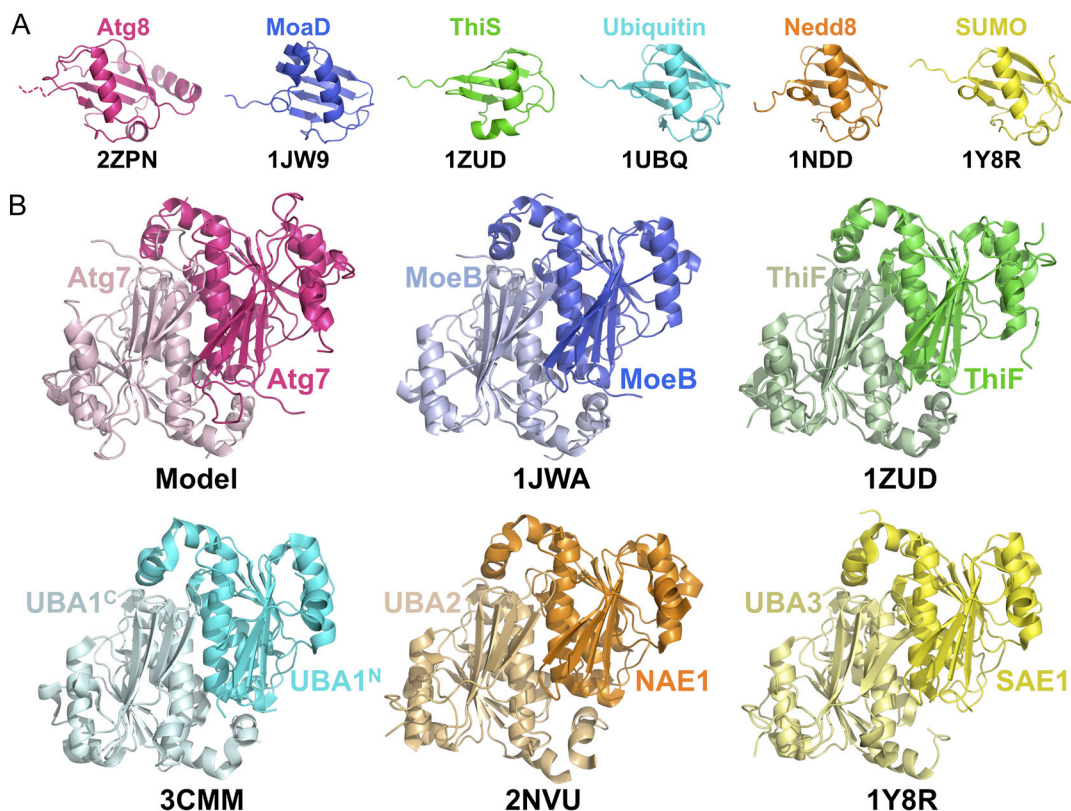


Figure 3-1. Structures of UBLs and their corresponding adenylation activation domains.

(A) Structures of UBLs displayed in the same orientation to show structural similarities. Atg8 is the only UBL that is not full-length (C-terminal tail of Atg8 depicted as dashed line). PDB codes are shown for each structure. (B) Structural model for dimeric Atg7 adenylation domain (generated using the Phyre server [126], details in materials and methods) with structures of corresponding adenylation domains from MoeB, ThiF, E1 for ubiquitin, E1 for Nedd8 and E1 for Sumo. The structural model and structures are depicted without their catalytic cysteine loop or domain for clarity. E1 for ubiquitin is composed of one polymer with the N-terminal half (UBA1^N) making up one “monomeric” adenylation domain with the C-terminal half (UBA1^C) making up the other. PDB codes are shown for each structure.

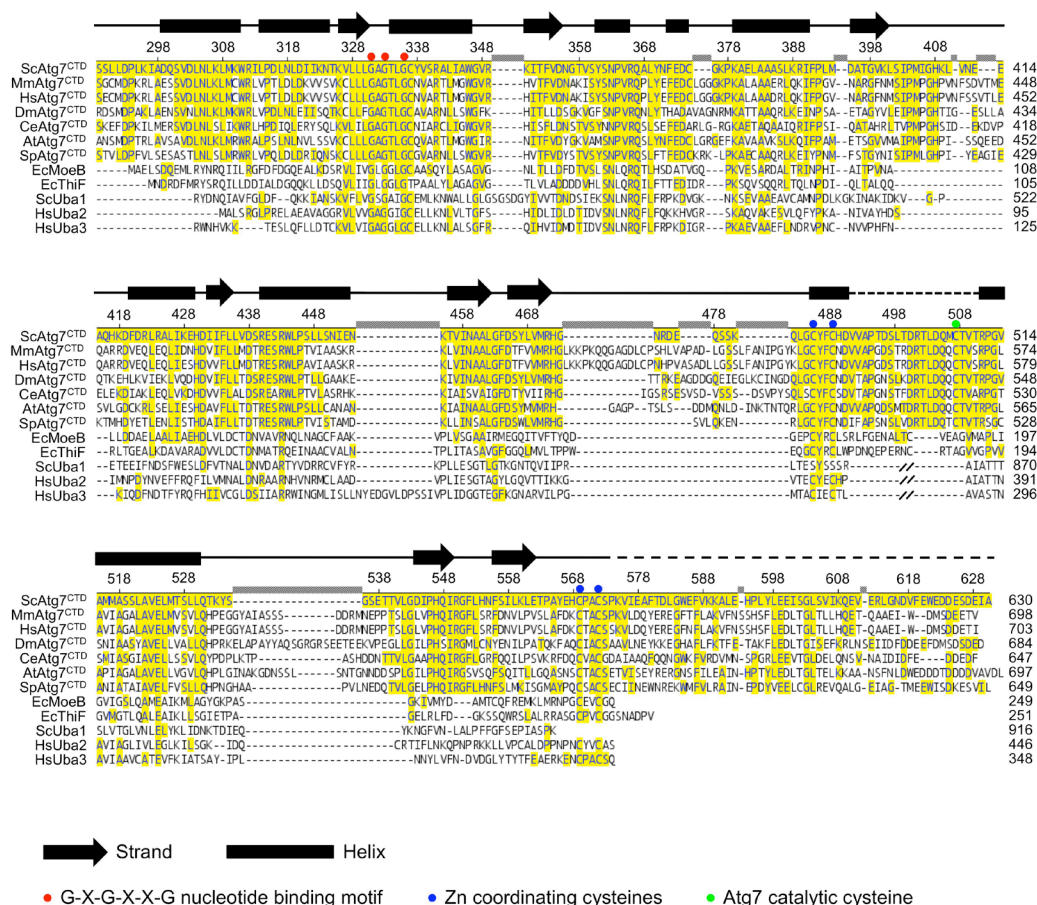


Figure 3-2. Atg7^{CTD} multiple sequence alignment.

Sequence alignment of *Saccharomyces cerevisiae* Atg7^{CTD} (Sc) with corresponding regions of Atg7 from mouse (Mm), human (Hs), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Arabidopsis thaliana* (At) and *Schizosaccharomyces pombe* (Sp), *Escherichia coli* MoeB (Ec), *Escherichia coli* ThiF, *Saccharomyces cerevisiae* Uba1, human Uba2 and human Uba3. Alignment is based on the structure of known crystal structures of MoeB, ThiF, Uba1, Uba2 and Uba3.

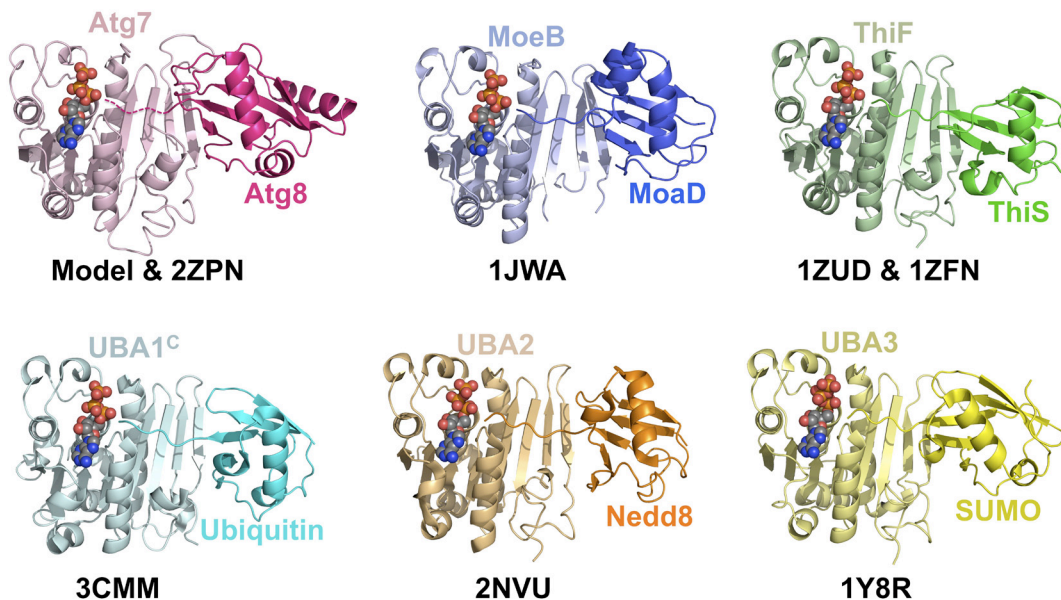


Figure 3-3. Monomeric adenylation domains bound to their corresponding UBL and a molecule of ATP.

Model of Atg7 adenylation domain (generated using the Phyre server [126], details in materials and methods) bound to Atg8 and ATP (spheres). The orientation of Atg8 was obtained by superimposition of the Atg7 model and Atg8 onto MoeB-MoaD (PDB code: 1JWA). In the case of the Atg7 model and UBA1^C, ATP was modeled in by superimposition with MoeB (PDB code: 1JWA). The structural model and structures are depicted without their catalytic cysteine loop or domain for clarity. ATP is shown as spheres in all cases.

two potential adenylation and thiolester active sites per Atg7 homodimer. By contrast, canonical E1s contain only a single adenylation active site, a single catalytic cysteine, and a single E2-binding site, all residing within the same subunit giving rise to structural pseudosymmetry with one “active” and one “inactive” half (reviewed in [3]). Furthermore, Atg7-Atg3 interactions differ from canonical E1-E2 interactions: Atg7 binds a unique Atg3 “flexible region” (referred to hereafter as Atg3^{FR}) that is not found in canonical E2s and is disordered in solution [78]. On the flip side, Atg7 also has unique features, most notably a ~300-residue N-terminal region that lacks sequence homology to any other proteins and whose function is unknown. In this study, we identify Atg7’s N-terminal domain (Atg7^{NTD}) as the binding site for the Atg3^{FR}. We find that Atg7 utilizes a unique mechanism that relies on its homodimerization to mediate E1-to-E2 UBL transfer *in trans*. Biophysical, biochemical, and X-ray crystallographic experiments reveal the structural basis and functional importance of distinctive E1-E2 interactions required for optimal Atg8 conjugation in the autophagy pathway.

Results and Discussion

Atg7 N-terminal Domain Recruits the Atg3 “Flexible Region”

To gain insights into the overall architecture of Atg7, we mapped two structural domains in the yeast protein based on prior studies [67, 77] and sequence analysis (Figures 3-2 and 3-4). Additionally cutting back into the first structural domain abolished protein expression. The Atg7^{NTD} spans residues 1 through 289 and does not display sequence homology to indicate structure or function. The Atg7 C-terminal domain (Atg7^{CTD}) sequence has two regions. First, residues 299-572 are homologous to the adenylation domains of canonical E1s for ubiquitin, NEDD8, and Sumo, as well as to the bacterial enzymes MoeB and ThiF, which catalyze C-terminal acyl-adenylation of MoaD and ThiS respectively. Both MoaD and ThiF have a UBL fold. The second region of Atg7^{CTD} is composed of ~60 C-terminal residues of Atg7 (Atg7^{C-ter}) that are required for expression in *E. coli* (no expression seen upon deletion of Atg7^{C-ter}) and presumably folding, and also for homooligomerization and UBL-binding in yeast [77]. A 9-residue linker (residues 290-298) connects the Atg7^{NTD} and Atg7^{CTD} (Figure 3-5).

Previous studies indicated that Atg7 forms a homodimer *in vivo* [77]. In order to understand the roles of our identified domains in assembly of exogenously expressed Atg7, we performed analytical ultra-centrifugation experiments on full-length yeast Atg7, Atg7^{NTD} and Atg7^{CTD}. The experimental molecular weights reveal that Atg7 and Atg7^{CTD} are dimeric in solution, whereas the Atg7^{NTD} is monomeric (Figures 3-5B-3-5D). Sedimentation equilibrium data indicate high-affinity dimer equilibrium dissociation constants for full-length Atg7 and Atg7^{CTD} of ~1 nM and ~30 nM respectively (Figures 3-5B and 3-5C).

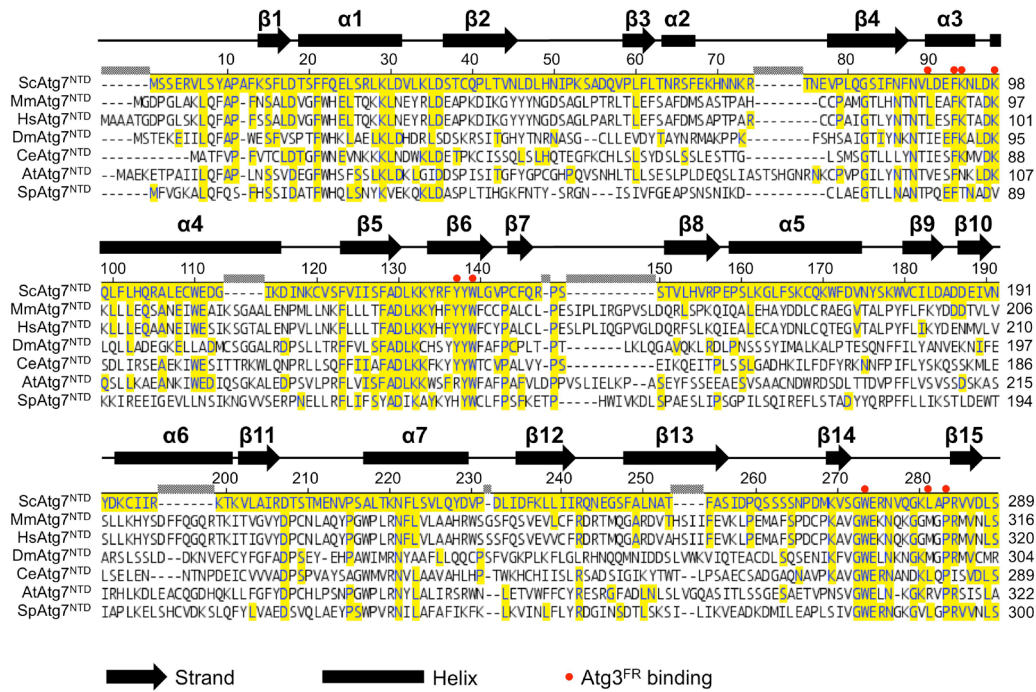


Figure 3-4. Atg7^{NTD} multiple sequence alignment.

Sequence alignment of *Saccharomyces cerevisiae* Atg7^{NTD} (Sc) with corresponding regions of Atg7 from mouse (Mm), human (Hs), *Drosophila melanogaster* (Dm), *Caenorhabditis elegans* (Ce), *Arabidopsis thaliana* (At) and *Schizosaccharomyces pombe* (Sp). Secondary structure assignment is based on the structure of Sc Atg7^{NTD} described herein (Figure 3-8).

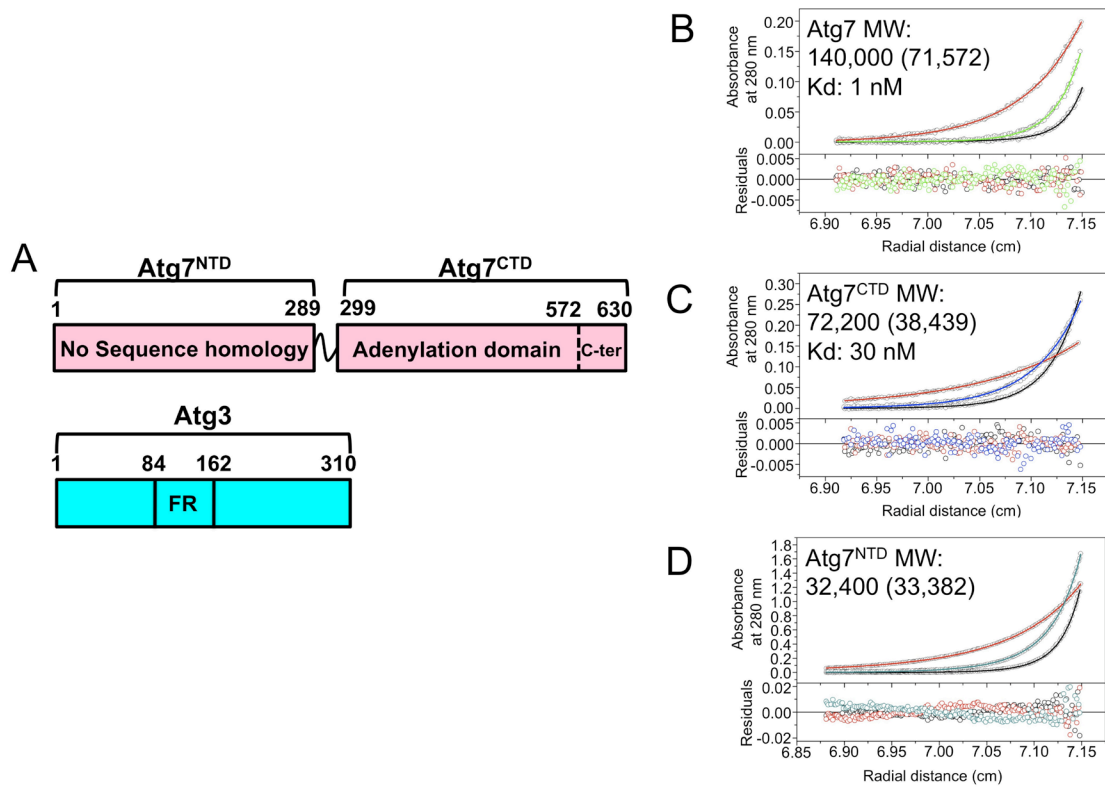


Figure 3-5. Atg7 is a homodimer, with the Atg7^{CTD} mediating the dimerization.

(A) Schematic of Atg7 and Atg3. (B) Sedimentation equilibrium AUC of Atg7. Samples between 0.86 and 3.5 μM were centrifuged at 4 °C at 12,000 (red), 18,000 (green) and 22,000 (black) rpm. The solid lines represent the global nonlinear least squares best-fit of the data. The molecular weight (MW) was determined using sedimentation velocity AUC. Theoretical MW is shown in parenthesis. (C) Sedimentation equilibrium AUC of Atg7^{CTD}. Samples between 0.86 and 3.5 μM were centrifuged at 4 °C at 12 k (red), 18 k (blue) and 22 k (black) rpm. The solid lines represent the global nonlinear least squares best-fit of the data to a reversible monomer-dimer self-association model with the shown dissociation equilibrium constant. The molecular weight (MW) was determined using sedimentation velocity AUC. Theoretical MW is shown in parenthesis. (D) Sedimentation equilibrium AUC of Atg7^{NTD}. Samples between 0.86 and 3.5 μM were centrifuged at 4 °C at 18 k (red), 27 k (cyan) and 34 k (black) rpm. The solid lines represent the global nonlinear least squares best-fit of the data. The molecular weight (MW) was determined using sedimentation velocity AUC. Theoretical MW is shown in parenthesis.

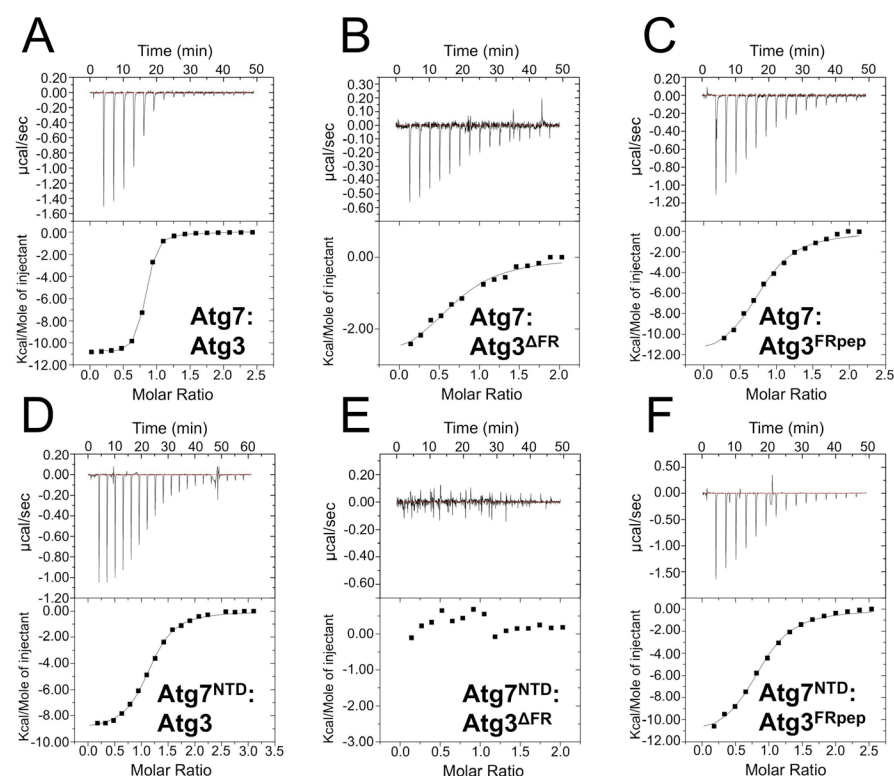
We next mapped high-affinity binding between Atg7 and Atg3, which interact with a K_d of 0.31 μM as measured by isothermal titration calorimetry (ITC) (Figure 3-6A). Atg7^{NTD} retains much affinity for Atg3, with a K_d of 1.83 μM (Figure 3-6D). A previous study showed that Atg3^{FR} (residues 84-161) binds Atg7, and an Atg3 mutant lacking the FR region (Atg3^{ΔFR}) was defective for autophagy [78]. Indeed, we found that deleting the FR region (residues 86-159) reduces the K_d for full-length Atg7 by a factor of ~60, and eliminates binding to the Atg7^{NTD} under our assay conditions (Figures 3-6B and 3-6E).

Given that Atg7 homodimerizes, we examined the ratios of Atg7:Atg3 interactions in the ITC experiments. Atg7^{NTD} binds Atg3 with approximately 1:1 stoichiometry (Figure 3-6G). The titration involving full-length Atg7 and Atg3 yields a stoichiometry of 0.77, which indicates that an Atg7 homodimer can bind two molecules of Atg3. Deviations from perfect 2:2 stoichiometry may arise from the presence of some aggregates that likely form during post-purification concentration necessary to obtain the 40-90 μM concentrations of full-length Atg7 required for the ITC experiments. Thus, overall, the ITC data suggest that the Atg7^{NTD}-Atg3^{FR} interaction constitutes a major, novel, stoichiometric E1-E2 interaction in the autophagy pathway.

To identify the region within the 80-residue Atg3^{FR} essential for binding Atg7^{NTD}, we examined co-pulldowns of His-MBP-tagged Atg3^{FR} deletion mutants and GST-tagged Atg7^{NTD} upon their coexpression in *E. coli*. Atg3 residues 130-142 appear to be both necessary and sufficient for binding to Atg7^{NTD} in this assay (Figure 3-7). Furthermore, a synthetic peptide encompassing Atg3 residues 128-144 (Atg3^{FRpep}) binds similarly to both full-length Atg7 and Atg7^{NTD}: 3.53 and 3.61 μM K_d , respectively (Figures 3-6C and 3-6F).

Structural Basis for Atg7^{NTD}-Atg3^{FRpep} Interactions

In order to gain detailed insights into Atg7-Atg3 interactions, we determined crystal structures of Atg7^{NTD} (Figures 3-8A and 3-9A), by single wavelength anomalous diffraction using selenomethionine substituted protein, and of Atg7^{NTD}-Atg3^{FRpep} complex (Figures 3-9A, 3-9D and 3-10), by molecular replacement using the structure of Atg7^{NTD} as a searchmodel (Table 3-1). Atg7^{NTD} adopts a novel fold, comprised of two structurally similar subdomains (Figure 3-8B). The subdomains are interconnected by three extended loops spanning the entire length of the Atg7^{NTD}, so the N- and C-termini are at opposite ends. Both subdomains consist of a 5-stranded mixed β -sheet sandwiched between two α -helices, with distinct flanking helices and loops. A search for homologous structures using the DALI server [108] identified several proteins having an MPN (Mpr1, Pad1 N-terminal) fold with Z-scores between 4 and 6: *Archaeoglobus fulgidus* AF2198 [127], human Mov34 [128], *Chlorobium tepidum* RadC (PDB code: 2QLC), yeast and *C. elegans* Prp8 [129], bacterial and eukaryotic cytidine deaminases [130, 131], and human AMSH-LP [132] (Figure 3-8C). MPN domains are thought to be protein-protein interaction domains, and also with the appropriate constellation of zinc ligands form



Macromolecule	Titrant	Kd (μM)	ΔH (cal/mol)	ΔS (cal/mol/deg)	N*
Atg7	ATG3	0.31	-1.09E+04	-6.25E+00	0.77
Atg7 ^{NTD}	ATG3	1.83	-9.14E+03	-3.89E+00	1.12
Atg7	ATG3 ^{ΔFR}	17.67	-3.12E+03	1.14E+01	0.70
Atg7 ^{NTD}	ATG3 ^{ΔFR}	-	-	-	-
Atg7	ATG3 ^{FRpep}	3.61	-1.26E+04	-1.66E+01	0.78
Atg7 ^{NTD}	ATG3 ^{FRpep}	3.53	-1.17E+04	-1.37E+01	0.84

Figure 3-6. Mapping the high-affinity binding site on Atg7 and Atg3 that mediates their interaction.

(A-F) ITC data for binding between Various Atg7 and Atg3 constructs. Upper panels show raw power data recorded during the titration experiments. Lower panels show fit of standard binding equations after integration of the raw data, using software provided by MicroCal. (G) Summary of thermodynamic parameters determined by ITC for data shown in panels E-J. * N value is shown for monomeric Atg7.

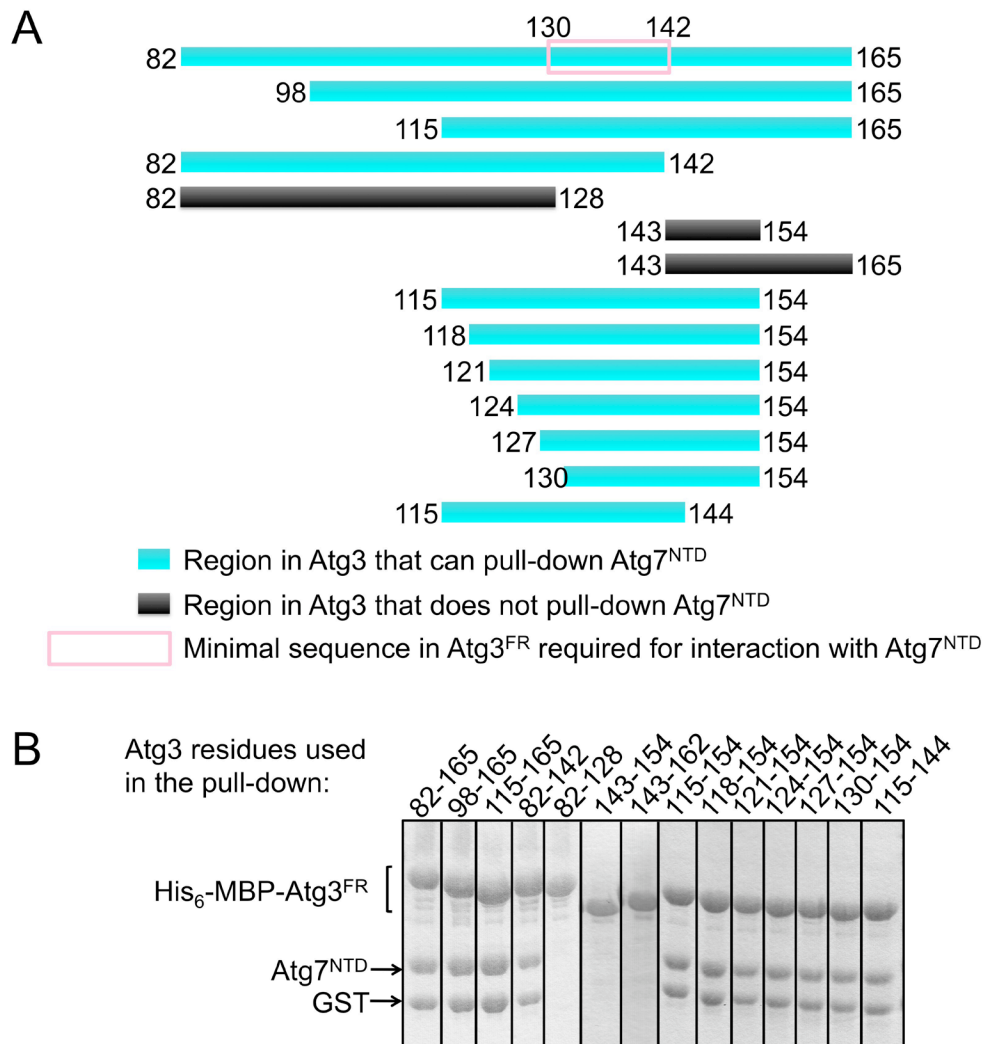


Figure 3-7. Mapping the minimal Atg3^{FR} region required for binding the Atg7^{NTD}.

(A) Schematic showing the various truncations of Atg3^{FR} used in pull-downs with Atg7^{NTD}. (B) Pull-downs from *E. coli* co-expressing wild-type Atg7^{NTD} and Atg3^{FR} fragments. Numbers above the gel panel correspond to the residues within Atg3 that were used in that pull-down.

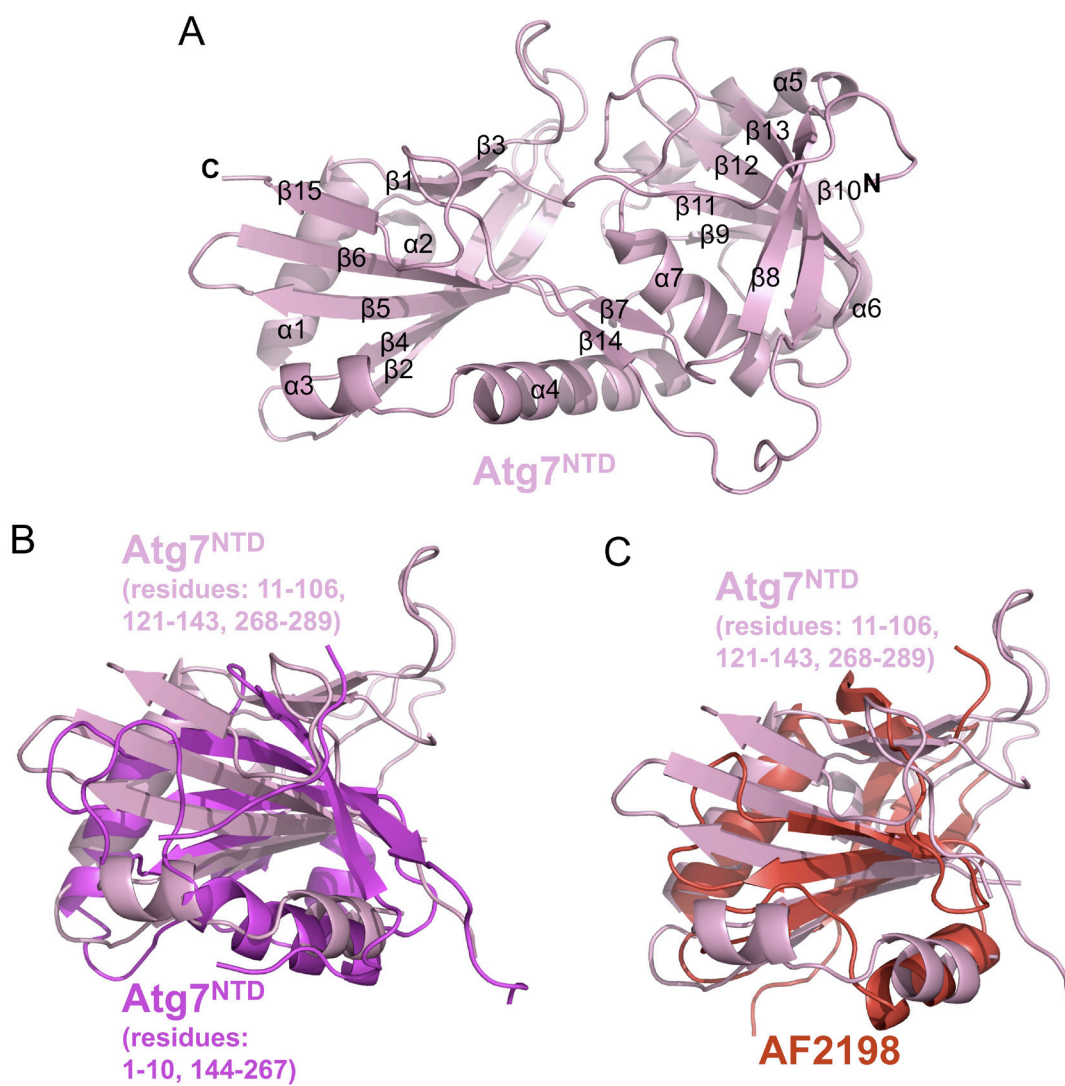


Figure 3-8. Structural features of Atg7^{NTD}.

(A) Structure of Atg7^{NTD} with strands and helices numbered. (B) Overlay of the two halves of Atg7^{NTD}. Residues 11-106, 121-143 and 268-289 from the Atg7^{NTD} structure (pink) overlaid with residues 1-10 and 144-267 from the Atg7^{NTD} structure (magenta). (C) Overlay of residues 11-106, 121-143 and 268-289 from the Atg7^{NTD} structure (pink) with the structure of *Archaeoglobus fulgidus* AF2198 (red, PDB code: 1OI0).

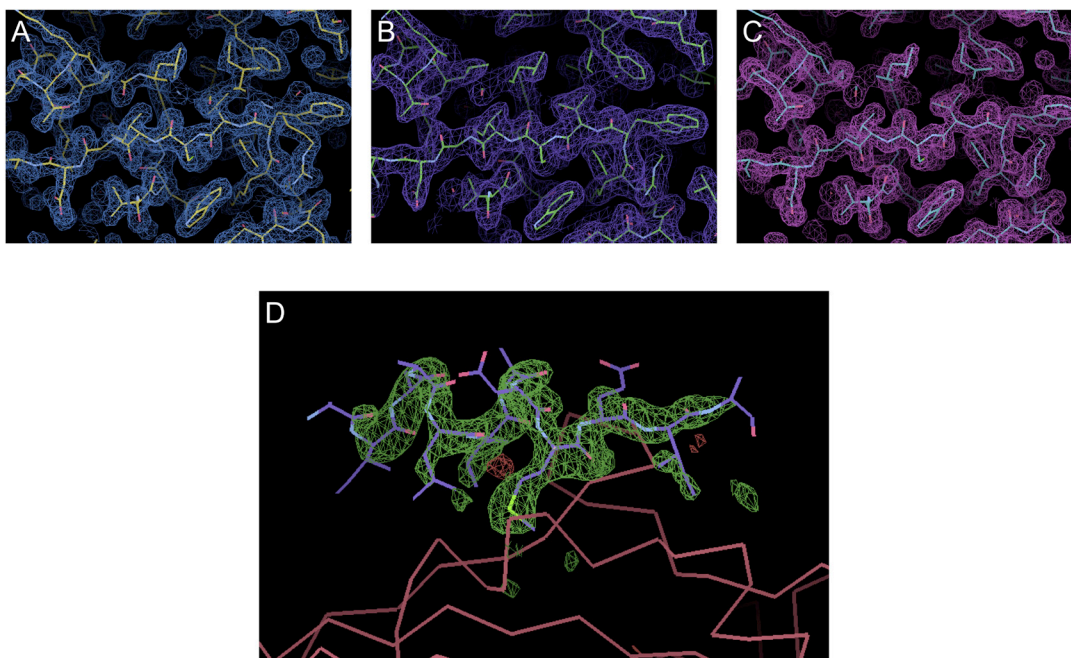


Figure 3-9. Final 2F_o-F_c electron density maps of Atg7^{NTD}.

(A) 2F_o-F_c electron density map contoured to 1σ (blue) over a region of Atg7^{NTD} (yellow). (B) 2F_o-F_c electron density map contoured to 1σ (purple) over a region of Atg7^{NTD} from the Atg7^{NTD}-Atg3^{FRpep} complex (green). (C) 2F_o-F_c electron density map contoured to 1σ (magenta) over a region of Atg7^{NTD}(P283D) (cyan). (D) F_o-F_c difference map contoured to 3σ (green) over the Atg3^{FRpep} (purple). The map was generated by performing simulated annealing starting at 4000 K, in 500 cooling steps to a final 300 K without the peptide present. Atg7^{NTD} displayed in ribbon (red).

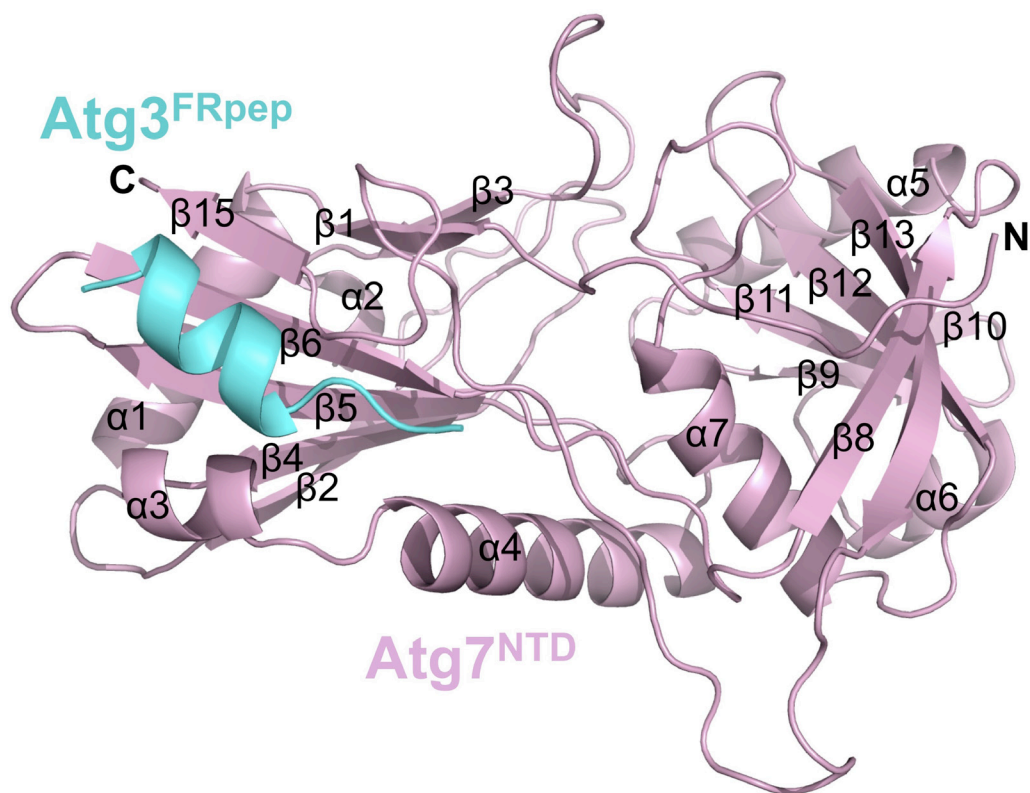


Figure 3-10. Structure of the Atg7^{NTD}-Atg3^{FRpep} complex.

Structure of the Atg7^{NTD}-Atg3^{FRpep} complex, with Atg7^{NTD} in pink and Atg3^{FRpep} in cyan. The Atg3^{FRpep} adopts a helical structure. Strands and helices are numbered on Atg7^{NTD}.

Table 3-1. Crystallographic and refinement statistics for Atg7^{NTD}, Atg7^{NTD}-Atg3^{FRpep} and Atg7^{NTD}(P283D).

Structure	Atg7 ^{NTD}	Atg7 ^{NTD} - Atg3 ^{FRpep}	Atg7 ^{NTD} (P283D)
Data collection			
Beamline	SERCAT ID24	ALS 822	SERCAT ID24
Wavelength (Å)	0.97924	1	1
Space group	P2 ₁ 2 ₁ 2	C2	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
a, b, c (Å)	a=56.856, b=74.677, c=76.047	a=152.135, b=99.429, c=71.502	a=56.184, b=74.439, c=153.373
α, β, γ (°)	α= β= γ=90 40.0-1.9 (1.97-	α= γ=90, β=113 50.0-2.1 (2.18-	α= β= γ=90 50-1.6 (1.66-
Resolution (Å)	1.90)	2.10)	1.60)
Total reflections	662452	721148	2400515
Unique reflections	26587	58224	85691
R _{merge} (%)	8.1 (25.0)	6.7 (51.2)	5.8 (45.0)
Average I/σ	43.3 (10.5)	22.4 (2.6)	31.5 (3.0)
Completeness (%)	97.3 (84.7)	99.9 (100.0)	99.8 (98.4)
Redundancy	13.3 (9.3)	3.7 (3.7)	7.0 (5.7)
Wilson B-factor	13.5	35.8	16.33
Refinement			
Resolution range (Å)	33.8-1.9	33.2-2.1	45.3-1.6
No. of reflections (σ≥0)	25621	55246	81152
R _{work} (%)	17.5	18.1	18.6
R _{free} (%)	19.7	20.9	20.6
Number of protein atoms	2343	4566	4636
Number of waters	323	193	689
Average B-factor (protein)	16.1	62.9	20.2
Average B-factor (water)	27.8	51.1	34.1
RMSD:			
Bond lengths (Å)	0.007	0.007	0.007
Bond angles (°)	1.025	1.012	1.036
Ramachandran plot statistics			
Residues in preferred regions (%)	97.8	95.0	97.1
Residues in additional allowed regions (%)	1.8	4.8	2.7
Residues in disallowed regions (%)	0.3	0.2	0.2

Table 3-1 (continued).

Highest resolution shell is shown in parenthesis. $R_{\text{work}} = \Sigma |F_o - F_c| / \Sigma F_o$. R_{free} is the cross-validation of R-factor, with 5% of the total reflections omitted during model refinement.

hydrolases such as cytidine deaminases and AMSH-LP in the JAMM-family deubiquitinating enzymes (DUBs) [127-129, 132]. However, we cannot speculate as to any functional significance of this structural homology, as Atg7^{NTD} lacks numerous features associated with zinc and UBL-binding by JAMM DUBs, and the UBL surface recognized by JAMM-family DUBs would presumably be sequestered upon Atg8 binding to the Atg7^{CTD}.

A groove in Atg7, which has β -strands-5, -6 and -15 as its floor and α -helix-3 and the β 14/ β 15-loop (not part of the core MPN-fold) as its sides, cradles the Atg3^{FRpep}. In complex with Atg7^{NTD}, the Atg3^{FRpep} adopts a 2-turn helix followed by a short extension (Figure 3-10). Interestingly, NMR data previously indicated in the absence of Atg7, the Atg3^{FR} adopts a disordered structure both on its own and in the context of full-length Atg3 [78]. Furthermore, it seems that the Atg3^{FRpep} can adopt a range of structures, because the corresponding sequence is the only portion of the FR visible in a prior crystal structure of full-length Atg3 (PDB code: 2DYT), where the Atg3^{FRpep} sequence forms a helix [78]. Consistent with its structural flexibility, in the crystal structure of full-length Atg3, this region has substantially higher B-factors than the rest of Atg3 (91 v. 51), half the side-chains including that of Met139 are not observed, and because the rest of the FR is not observed it is possible that the Atg3^{FRpep} sequence comes from an adjacent Atg3 in the crystal to mediate packing (Figure 3-11). It seems likely that the structural malleability of the Atg3^{FRpep} region facilitates accessibility of this region for binding to Atg7.

The Atg7-Atg3^{FRpep} interaction is mediated by hydrophobic side-chains from Atg3 inserting into a hydrophobic groove in Atg7 (Figure 3-12). The interface is centered around Atg3 residue Met139 docking in a hydrophobic pocket formed by Atg7 residues Phe93, Lys94, Trp139, and Pro283. This is buttressed on one side by Atg3 residue Leu135 making hydrophobic contacts with Atg7 residues Leu90, Phe93, Lys94, and Tyr137. On the other side, Atg3 residue Ile141 interacts with Atg7 residues Lys98, Trp139, Trp273, and Leu281. In total, the overall interaction buries $\sim 940 \text{ \AA}^2$ of surface area.

Role for Atg7^{NTD}-Atg3^{FRpep} Interface in Atg7-Atg3 Interactions and in the Transfer of Atg8

We tested whether the structurally observed interface is important for the Atg7^{NTD}-Atg3 interaction using our pull-down assay after coexpression of wild-type (WT) and mutant versions of GST-tagged Atg7^{NTD} and His-MBP-tagged Atg3. Because of the hydrophobic nature of the interaction, we tested the effects of pairs of hydrophobic to Ala mutations in Atg3 to remove interacting side-chains. From Atg7, we made single hydrophobic to Asp mutations, to help alter the overall hydrophobic character of the groove. Pull-down performed with the His-MBP tag shows that WT Atg3 is able to pull down similar amounts of WT Atg7^{NTD}. Pairwise Ala mutations in Atg3 Leu135, Met139, and Ile141, or individual Asp substitutions in place of Atg7^{NTD} Tyr137, Trp139, or Pro283 eliminated the interaction (Figure 3-13A). In the reciprocal pull-down for GST-

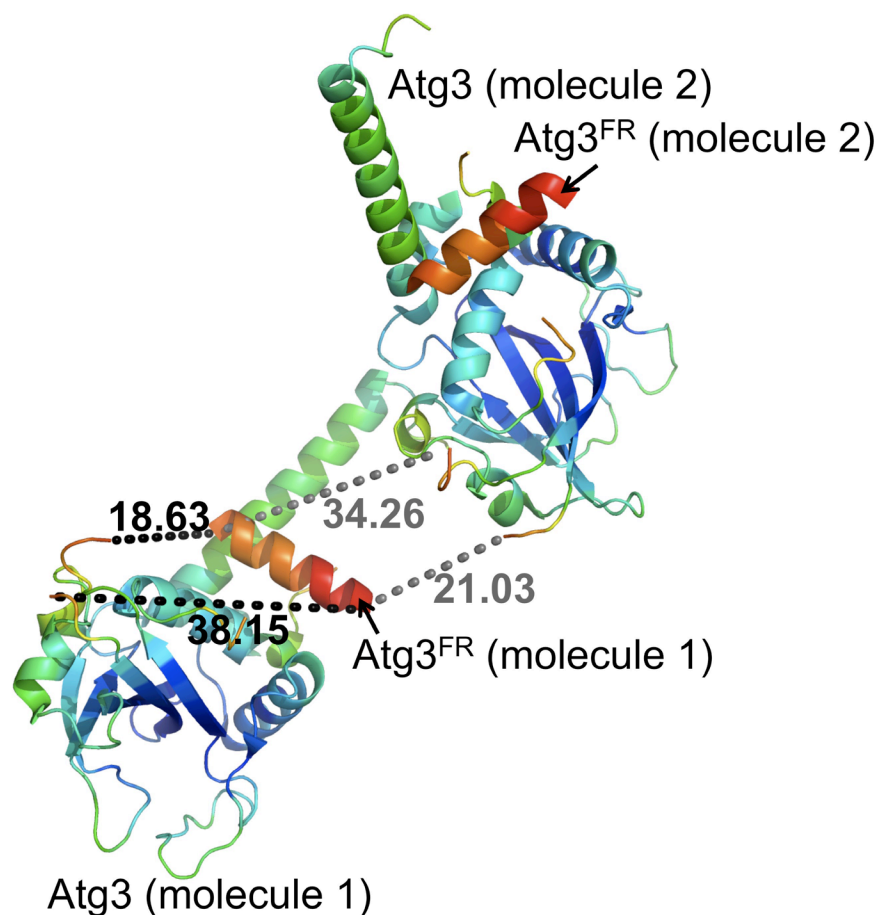


Figure 3-11. The FR region of Atg3 is flexible.

Two molecules of Atg3 (PDB code: 2DYT) generated by symmetry are colored by C- α B-factors, with regions in red having the highest B-factors and regions in blue having the lowest B-factors. Distance between Atg3^{FR} of molecule 1 with the core Atg3 of molecule 1 is shown in black (total measured distance is 56.78Å). Distance between Atg3^{FR} of molecule 1 with the core Atg3 of molecule 2 is shown in grey (total measured distance is 55.29Å).

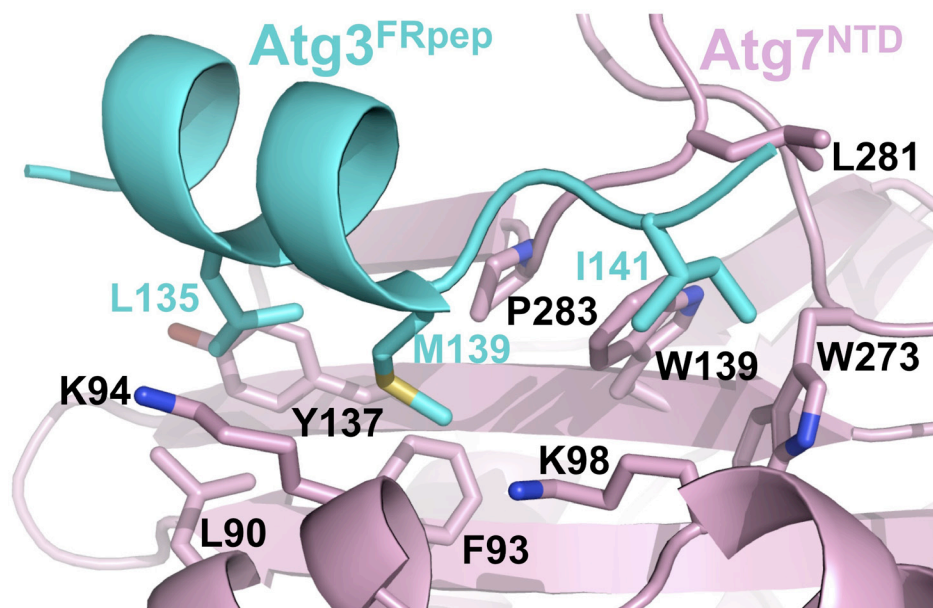
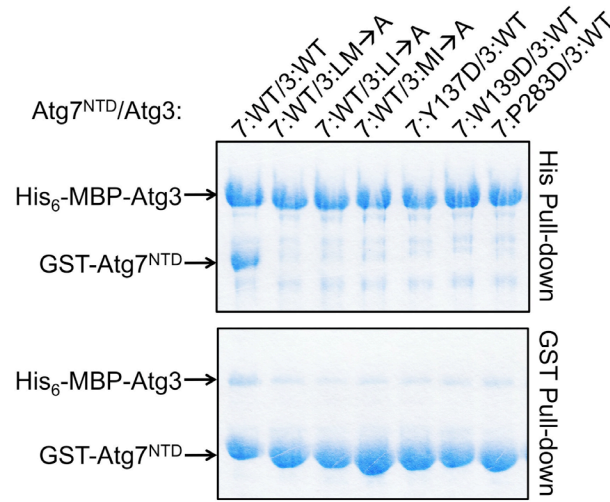


Figure 3-12. The Atg3^{FRpep} mediates interactions with Atg7^{NTD} via a hydrophobic interface.

Residues involved in the interaction shown in sticks. Atg7^{NTD} is colored pink and Atg3^{FRpep} is colored cyan. Oxygen atoms are colored red, nitrogen atoms are colored blue and sulfur atoms are colored yellow.

A



B

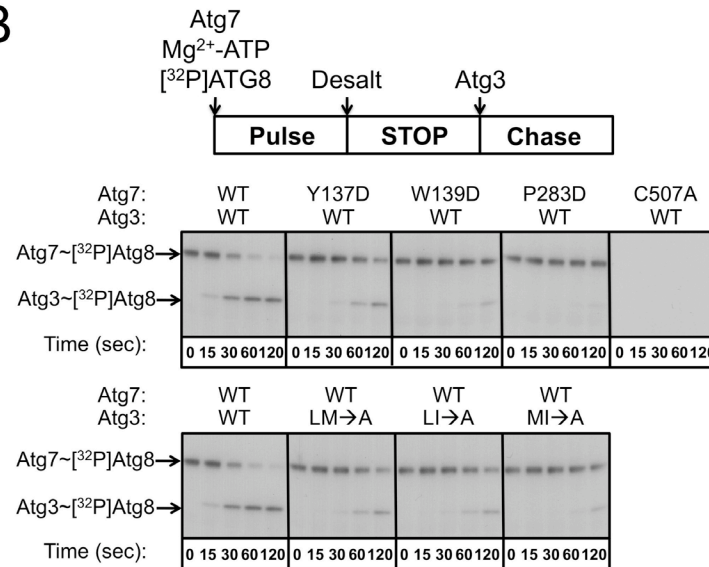


Figure 3-13. Atg7^{NTD}-Atg3^{FRpep} interface is important for Atg7-Atg3 interactions and for the transfer of Atg8.

(A) Pull-downs from *E. coli* co-expressing wild-type and mutant versions of Atg7^{NTD} and Atg3^{FR}. Mutations are at the binding interface of the two proteins. WT denotes wild-type protein; LM→A denotes Atg3(L135A, M139A); LI→A denotes Atg3(L135A, I141A); and MI→A denotes Atg3(M139A, I141A). (B) Top: schematic summarizing how the pulse-chase enzyme assay was performed. Bottom: autoradiograms showing time course of the transfer of [³²P]Atg8 from Atg7 to Atg3 (transthiolation reaction). WT denotes wild-type protein; LM→A denotes Atg3(L135A, M139A); LI→A denotes Atg3(L135A, I141A); and MI→A denotes Atg3(M139A, I141A).

Atg7^{NTD}, even for the WT proteins, only a substoichiometric amount of His-MBP-Atg3 copurifies, as would be the case if GST-Atg7^{NTD} were present in significant excess during the coexpression. Nonetheless, reduced interaction is observed for all the mutants, all of which are expressed at similar levels (Figure 3-13A).

We next tested whether the structurally observed Atg7^{NTD}-Atg3^{FRpep} interaction plays a role in Atg8 transfer from Atg7 to Atg3 using a pulse-chase assay. Briefly, in the “pulse” reaction we incubated Atg7 with [³²P]Atg8 and Mg²⁺-ATP to form the thiolester linked Atg7~[³²P]Atg8 intermediate, and stopped this reaction by desalting into buffer without Mg²⁺-ATP. We added Atg3, and monitored transfer of [³²P]Atg8 from Atg7 onto Atg3, yielding the thiolester-linked Atg3~[³²P]Atg8 conjugate (Figure 3-13B). In agreement with the structure, mutations at the Atg7^{NTD}-Atg3^{FRpep} interface cripple transfer of [³²P]Atg8 from Atg7 to Atg3 (Figure 3-13B).

In order to provide support for the notion that the deleterious effects of Atg7^{NTD} mutations were from eliminating the interaction with Atg3^{FR} and not due to global misfolding, we determined the crystal structure of the most impaired mutant, which has the Pro283 to Asp substitution. We screened for new crystallization conditions, because both Pro283 and a selenomethionine (residue 1) participate in crystal contacts for apo-Atg7^{NTD} but are not present in the methionine-containing mutant protein used for crystallization. Despite subtle differences in crystal packing localized around these residues, the Atg7^{NTD} and Pro283Asp mutant structures superimpose with 0.55 Å root-mean-square deviation (Figures 3-9C and 3-14).

The Atg7^{NTD} Binding Surface for Atg3^{FRpep} is Conserved Across Species

Sequence analysis of the Atg7^{NTD} suggests that its interaction with the Atg3^{FR} is conserved across eukaryotes (Figures 3-4 and 3-15) [133]. Indeed, as with the proteins from budding yeast, after coexpression in *E. coli* we observed interaction between GST-tagged mouse Atg7^{NTD} (spanning residues 1-317 of mouse Atg7) and either His-MBP-tagged mouse Atg3 or mouse Atg3^{FR} (residues 91-193 of mouse Atg3). Individual Asp mutations in place of mouse Atg7^{NTD} Tyr141, Trp143, and Pro310, which correspond to yeast Tyr137, Trp139, and Pro283, respectively, diminish the interaction (Figure 3-16). To test a function *in vivo*, we stably expressed an N-terminally GFP-tagged mouse Atg7, or the Pro310Asp mutant, or a catalytic Cys507Ala mutant, in Atg7 null MEFs. Chloroquine was added to induce accumulation of autophagosomes. Notably, accumulation of LC3-II was impaired by the mutation in the ATG3^{FRpep} binding site (Figure 3-17). The deleterious effects of the mutation are specific for Atg3, because no effect is observed for conjugation of Atg12 onto Atg5 (Figure 3-17), which is mediated by a different E2, Atg10, that is homologous to much of the Atg3 core domain but lacks an obvious FR sequence.

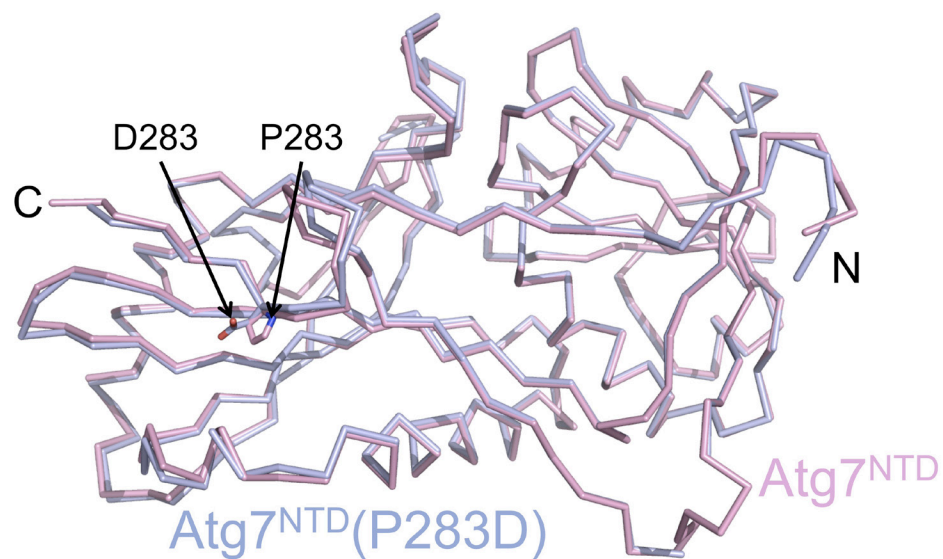


Figure 3-14. Overlay of the crystal structure of Atg7^{NTD} with that of Atg7^{NTD}(P283D).

Atg7^{NTD} (pink) and Atg7^{NTD}(P283D) (blue) superimpose with a 0.55 Å root-mean-square deviation.

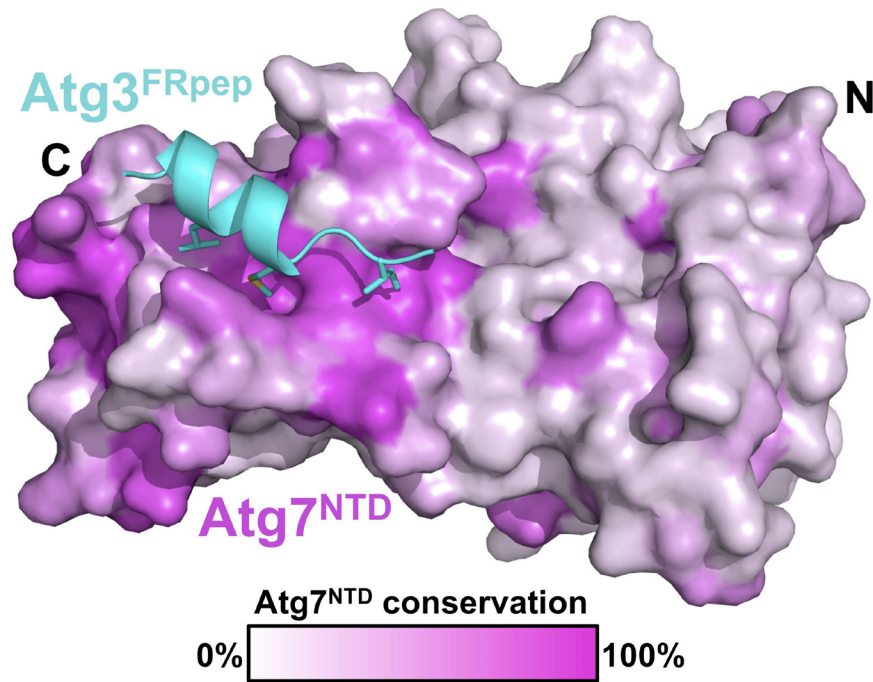


Figure 3-15. Atg3^{FR} binds a conserved surface on Atg7^{NTD}.

Conservation among Atg7^{NTD}s displayed as surface representation in the Atg7^{NTD}-Atg3^{FR}pep complex structure. White represents 0 % conservation while pink represents 100 % conservation across species. Leu135, Met139 and Ile141 on Atg3^{FR}pep shown in sticks.

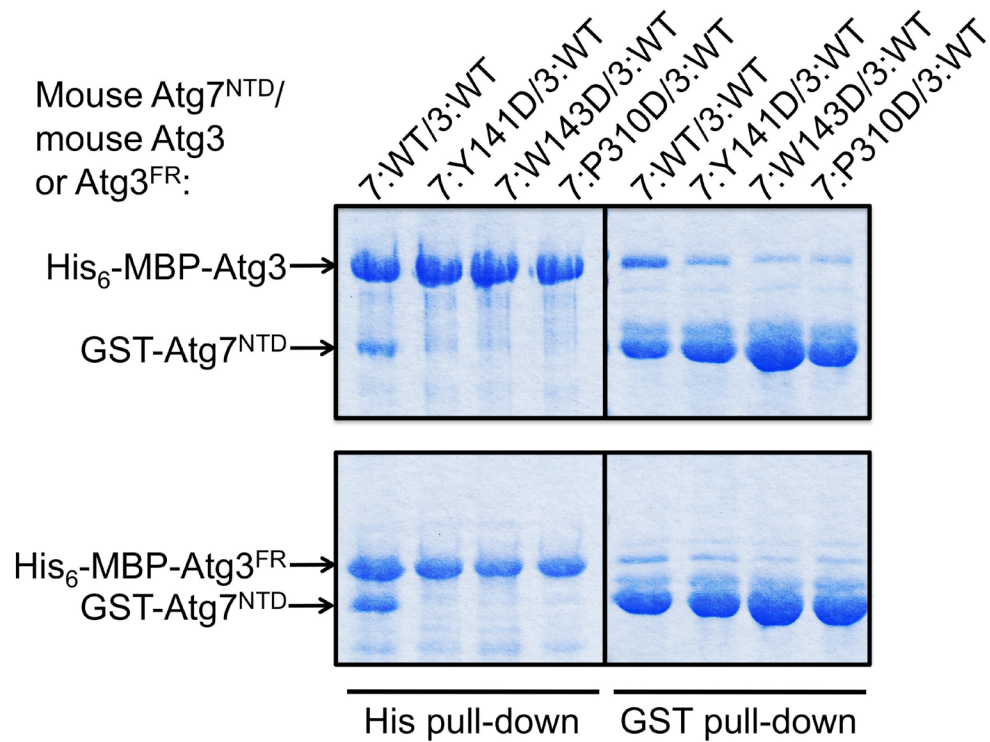


Figure 3-16. Mouse Atg7 uses the same conserved surface as its yeast counterpart to bind mouse Atg3^{FR}.

Pull-downs from *E. coli* co-expressing wild-type (WT) and mutant versions of mouse Atg7^{NTD} and WT mouse Atg3 or Atg3^{FR}. Mutations in mouse Atg7^{NTD} are based on sequence conservation at its Atg3^{FR} binding site.

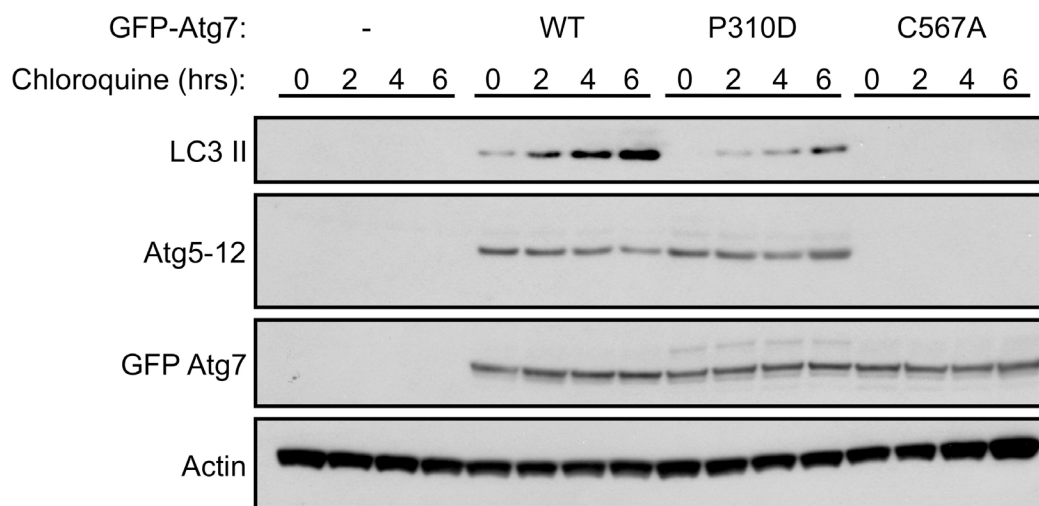


Figure 3-17. Mutation on Atg7's Atg3^{FR} binding surface leads to defects in autophagy.

Atg7 null MEF cell lines were made to stably express GFP-Atg7 WT and mutant versions. Chloroquine was added to prevent autophagic flux. Westerns performed on samples collected over a time course using primary antibodies against LC3, Atg12, Atg7 and actin.

Model for Transfer of Atg8 from Atg7 onto Atg3: a Unique Trans Mechanism Requiring Optimal Spacing between the Atg7^{NTD} and Atg7^{CTD}

Given the importance of the Atg7^{NTD}-Atg3^{FRpep} interface, we wished to make a structural model that would give insights into the roles of this interaction in the context of full-length Atg7 and Atg3. Briefly, we first obtained a homology model for the majority of the Atg7^{CTD}, including the catalytic Cys, using the Phyre server [126] based on sequence similarity to *E. coli* MoeB (Figure 3-2). We superimposed two copies of this model onto the MoeB dimer to obtain a dimeric Atg7^{CTD} lacking only the C-ter domain. We next placed the Atg7 and Atg3 catalytic cysteines near each other, and modeled the Atg7^{NTD}-Atg3^{FRpep} within reasonable distances to their connections in the remainder of Atg7 and Atg3, respectively. The model suggests a mechanism of E1-E2 transthiolation that differs from canonical E1s and E2s such as for ubiquitin, NEDD8, and Sumo, in which the same E1 subunit contains both the catalytic Cys and E2-binding regions, for UBL transfer in cis. A novel “trans” mechanism provides a rationale for Atg7 homodimerization: Atg8 linked to the catalytic Cys of one Atg7^{CTD} would be transferred to Atg3 associated with the Atg7^{NTD} of the opposite Atg7 protomer in the dimer (Figures 3-18 and 3-19).

To test the model, we coexpressed His- and GST-tagged Atg7 in *E. coli*, and sequentially performed affinity capture with Ni²⁺ agarose beads and then with glutathione sepharose beads to obtain Atg7 dimers containing two different protomers. Using this method, we made two mutant Atg7s. Atg7^{trans} contains two mutant versions of Atg7 that together form a dimer: one mutant has an Ala substitution in place of the catalytic Cys507, and the other has the Pro283Asp mutation that hinders the Atg7^{NTD} from binding Atg3. Our sequential affinity capture purification yields a mixed dimer where one protomer is unable to form a thiolester with Atg8 but retains Atg3 FR binding, and the other protomer is able to form a thiolester linkage with Atg8, but is defective in recruiting the Atg3 FR (Figures 3-20A and 3-21A). Therefore, the only way for Atg7^{trans} to mediate Atg8 transfer would be if Atg3 binds one Atg7^{NTD} and accepts Atg8 from the catalytic Cys of the opposite Atg7 in the dimer. By contrast, the Atg7^{cis} dimer contains one wild-type Atg7 and one double Cys507Ala/Pro283Asp mutant that would be defective in both forming a thiolester with Atg8 and binding to Atg3 (Figures 3-20A and 3-21B). Thus, Atg7^{cis} would only be active if Atg8 were transferred between the Atg7 catalytic Cys and an Atg3 associated with the Atg7^{NTD} of the same protomer of Atg7. Using our pulse-chase assay, we find that Atg7 preferentially carries out transthiolation in trans (Figure 3-20B).

We wondered whether the flexibility and relative placement of the Atg7^{NTD}, which binds the Atg3^{FR}, and the Atg7^{CTD}, whose catalytic Cys is thiolester-linked to Atg8, influences Atg8 transfer between Atg7 and Atg3. Thus, we made a series of mutations in the linker between the domains: a Gly mutation in place of the conserved Pro294 would test the effects of increasing linker flexibility while maintaining the linker length, a deletion mutant lacking residues 290-294 would test the effects of shortening the linker, and a mutant with the sequence Gly-Gly-Ser-Gly inserted in between residues Ser290 and Leu291 would test the effects of lengthening the linker. The Pro294Gly

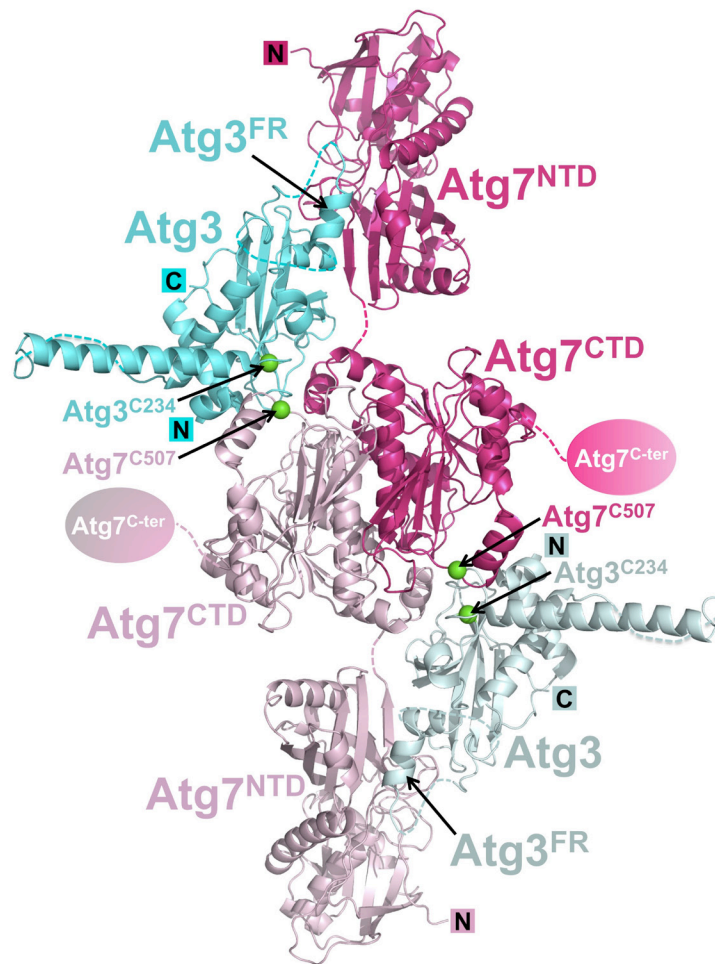


Figure 3-18. Transthiolation in trans: a working model for transfer of Atg8 from Atg7 to Atg3.

(A) Structural model of a homodimeric Atg7 (dark and light pink) bound to two Atg3 molecules (cyan and light cyan) in a catalytically active conformation, with the catalytic cysteines of each Atg3 (C234) poised to attack Atg7's catalytic cysteines (C507); catalytic cysteines shown as green spheres. The ~60 C-terminal residues (Atg7^{C-ter}) of each Atg7 homodimer is represented as an oval. For clarity, Atg8, which would start out thioester bound to the catalytic cysteines of Atg7 and end ligated to the catalytic cysteine of Atg3, is not shown.

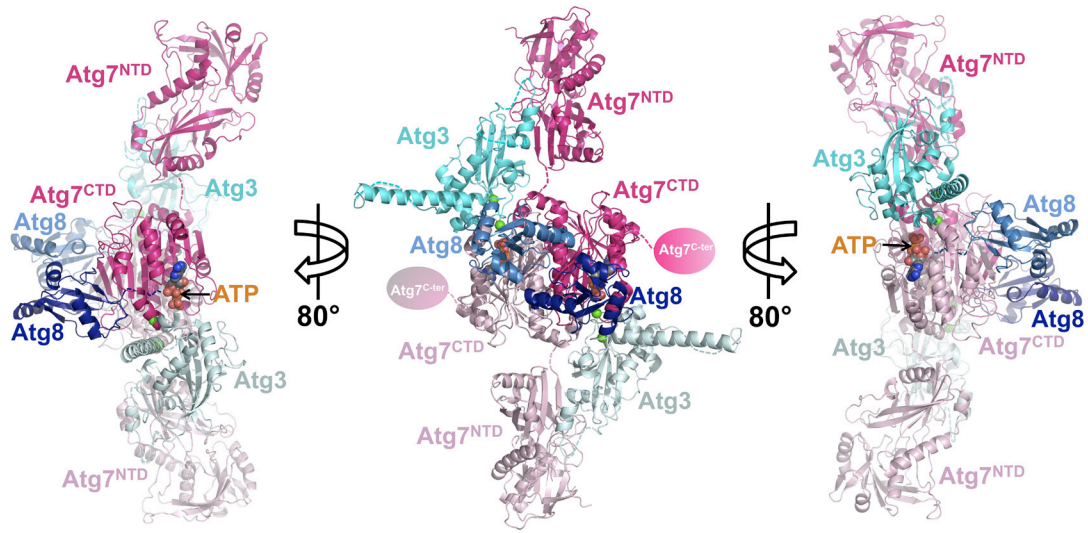


Figure 3-19. Transthiolation model of dimeric Atg7 bound to two molecules of Atg3, Atg8 and ATP.

Model generated as described in material and methods. Model is identical to that in Figure 3-13A with the addition of Atg8 (blue) and ATP (spheres) bound to the adenylation domain. Atg8's C-terminal tail is missing from the structure and is drawn as dashed lines. Catalytic cysteines of Atg7 and Atg3 shown as green spheres.

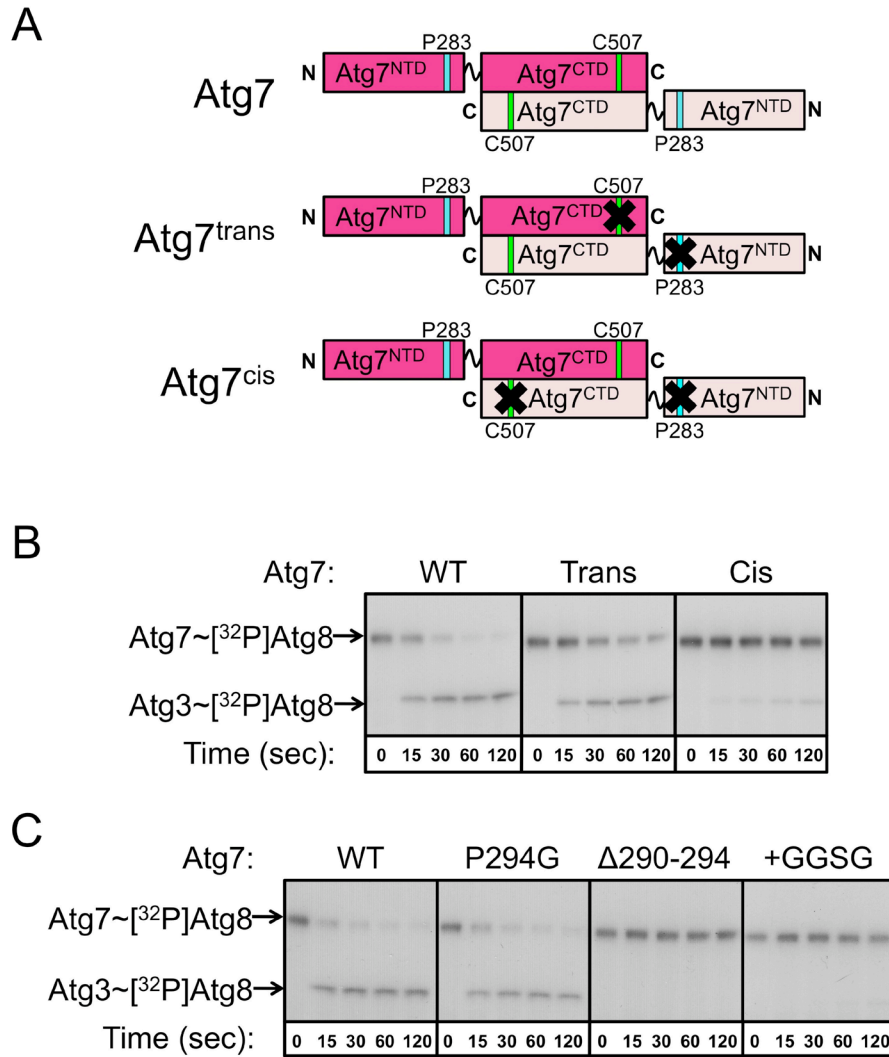


Figure 3-20. Atg7 performs the transthiolation reaction in trans and has a requirement for its linker length between the Atg7^{NTD} and Atg7^{CTD}.

(A) Schematic explaining Atg7^{trans} and Atg7^{cis}. Residue Pro283 that binds Atg3^{FR} is represented in cyan and Atg7's catalytic cysteine (Cys507) is shown in green. The black X indicates mutation of Pro283 to Asp or Cys507 to Ala that impairs binding of Atg3 or prevents thiolester formation with Atg8 respectively (B) Autoradiogram showing time course of the transfer of [³²P]Atg8 from Atg7 to Atg3 (transthiolation reaction). WT denotes wild-type protein; Trans denotes an Atg7 homodimer composed of Atg7(P283D) and Atg7(C507A); Cis denotes an Atg7 homodimer composed of wild-type Atg7 and Atg7(P183D, C507A). (C) Autoradiogram showing time course of the transfer of [³²P]Atg8 from Atg7 to Atg3 (transthiolation reaction). WT denotes wild-type protein; +GGSG denotes an insertion of GGSG between position 290 and 291 in Atg7.

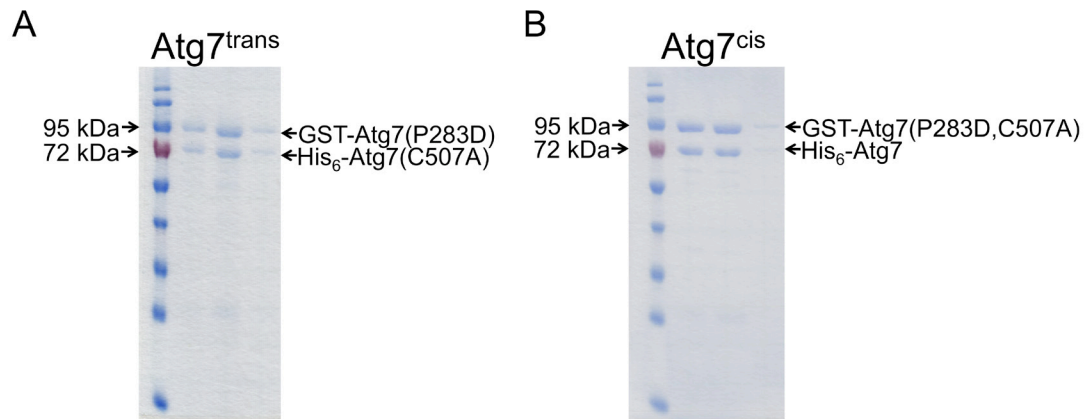


Figure 3-21. Gels showing Atg7^{trans} and Atg7^{cis} during purification.

(A) Gel showing Atg7^{trans} after initial double purification using Ni²⁺ agarose and glutathione sepharose affinity chromatography, before cleavage of GST and His₆ tags.
 (B) Gel showing Atg7^{cis} after initial double purification using Ni²⁺ agarose and glutathione sepharose affinity chromatography, before cleavage of GST and His₆ tags.

mutant displayed wild-type ability to mediate transfer of Atg8 to Atg3, whereas the deletion and insertion mutants were severely impaired (Figure 3-20C). The results are consistent with the notion that the Atg7^{NTD} and Atg7^{CTD} are flexibly tethered, but that the distance between them is optimized for the transthiolation reaction.

Implications for E1~UBL-E2 Cascades in Autophagy

We find that Atg7's unique N-terminal domain selectively recruits Atg3's FR via a novel type of E1-E2 interaction. First, Atg7^{NTD} forms a roughly stoichiometric complex with Atg3, which is mediated by the Atg3^{FR}. Second, the crystal structure of yeast Atg7^{NTD} reveals a novel fold harboring a conserved hydrophobic groove recruiting Atg3^{FR}. Third, on this basis, we were able to introduce a mutation into mouse Atg7 that leads to defects in conjugation of LC3, the mammalian ortholog of yeast Atg8. Finally, we identify a novel "trans" E1-to-E2 Atg8 transfer mechanism that provides a rationale for the homodimerization of Atg7: Atg3 bound via its FR to one Atg7^{NTD} accepts Atg8 from the catalytic Cys of the opposite Atg7 protomer in the dimer.

In addition to Atg7 interacting with Atg3 during conjugation of the UBL Atg8, Atg7 has the additional function of activating a second UBL, Atg12, which functions via the distinct E2, Atg10. Notably, we have been able to introduce a mutation in mouse Atg7 that separates these two pathways. The Pro310Asp mutation in the mouse Atg7 NTD, made based on our crystal structure of yeast Atg7^{NTD}-Atg3^{FR}, both hinders binding

to Atg3, and LC3 conjugation in cells. This had no effect on Atg12 conjugation in cells (Figure 3-17). Consistent with this notion, Atg10 sequences lack an obvious FR. Taken together with our crystallographic data indicating that Atg3^{FRpep} binds in a groove in Atg7, the results raise the possibility that it may be possible to obtain small molecule inhibitors that selectively inhibit Atg8 conjugation.

How might Atg7 interact with its distinct E2s? Although our data do not address the mechanism of Atg12 conjugation, we speculate that the answer to this question is related to our finding that Atg3 binds full-length Atg7 with a 6-fold greater affinity than it binds Atg7^{NTD}. Thus, in addition to the high-affinity Atg7^{NTD}-Atg3^{FR} interaction investigated here, it seems likely that there are additional weak contacts between the Atg7^{CTD} and Atg3^{ΔFR}. We speculate that these are likely mirrored by Atg10, which displays sequence homology to Atg3^{ΔFR}. Moreover, in our structural model of an Atg7-Atg8-Atg3 complex, Atg3 is located near Atg8, raising the possibility that the UBLs Atg8 and Atg12 help mediate specificity of transfer to their own distinct E2s. Future studies will be required to understand the detailed molecular basis for these other interactions and their roles in E1-to-E2 UBL transfer in autophagy.

Materials and Methods

Constructs, Protein Expression and Purification

All constructs were generated by standard molecular techniques and coding regions were sequence verified by the Hartwell Center for Bioinformatics and Biotechnology at St Jude Children's Research Hospital. Point mutants were generated using QuikChange (Stratagene). N-terminal GST-, His₆-MBP- or His₆-tag fusions were expressed in *E. coli* from regular or modified versions of pGEX4T1 (GE lifesciences), pGEX2Tk (GE lifesciences), pRSF1b (Novagen) or pRSF-Duet (Novagen) vectors. Protein tags were designed to be cleavable by either Thrombin or TEV protease and this leaves 2 extra residues (Gly-Ser) on the N-termini of proteins. Details of yeast constructs are as follows: Atg7 refers to residues 1-630 from Atg7, Atg7^{NTD} refers to residues 1-289 from Atg7, Atg7^{CTD} refers to residues 289-630 from Atg7, Atg3 refers to residues 1-310 from Atg3, Atg3^{FR} refers to residues 82-165 from Atg3, Atg3^{FRpep} refers to residues 128-144 from Atg3 and Atg3^{ΔFR} refers to deletion of residues 86-159 from Atg3. Details of mouse constructs are as follows: Atg7^{NTD} refers to residues 1-317 from Atg7. Atg3 refers to residues 1-314 from Atg3 and Atg3^{FR} refers to residues 91-193 from Atg3.

Retroviral constructs expressing mouse Atg7 and mutants were generated in multiple steps. First Atg7 (and mutants) were subcloned into the BamHI and EcoRI sites of retroviral expression vector LZRS-IRES-ZEO [134]. Second, eGFP (Clonotech) was cloned inframe, into the BamHI site resulting in the expression of eGFP-Atg7 fusion with a 2-residue flexible linker (sequence Gly-Ser).

All protein expression and purification for crystallography, AUC, ITC and enzyme assays was done in *E.coli*, namely BL21(DE3) Gold (Novagen), or codon enhanced RIL (Stratagene) or Rosetta (Novagen) cells grown in LB broth Miller (EMD) supplemented with the appropriate antibiotic. Protein purification protocols vary for different experiments and are described under the respective sections below.

Analytical Ultracentrifugation (AUC)

24 L of wild-type GST-Atg7 was grown in BL21(DE3) Gold at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 23°C. The cells were spun down and resuspended in buffer containing 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM DTT and 2.5 mM PMSF. Cells were then lysed by a combination of sonication and the use of an EmulsiFlex-C5 cell homogenizer (Avestin). Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. Glutathione affinity chromatography was used as the first step in the protein purification protocol using 20 ml of glutathione sepharose beads. After passing the supernatant over the beads, the beads were washed with 300 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT. Following elution of the protein with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione, the GST tag was cleaved using Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) overnight at 4°C. The cleaved protein was diluted to reduce the concentration of NaCl to ~40 mM using buffer containing 25 mM Tris pH 7.6 and 5 mM DTT. The diluted protein was subjected to Source Q anion exchange chromatography using an AKTA FPLC system using buffer containing 25 mM Tris pH 7.6, 5 mM DTT with NaCl gradient. Finally, the protein was applied to a Superdex-200 size-exclusion chromatography column using the AKTA FPLC system. The protein was concentrated and stored at -80°C in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT.

24 L of wild-type GST-Atg7^{CTD} was grown in BL21(DE3) Gold at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 16°C. The cells were spun down and resuspended in buffer containing 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM DTT and 2.5 mM PMSF. Cells were then lysed by a combination of sonication and the use of an EmulsiFlex-C5 cell homogenizer (Avestin). Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. Glutathione affinity chromatography was used as the first step in the protein purification protocol using 20 ml of glutathione sepharose beads. After passing the supernatant over the beads, the beads were washed with 300 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT. Following elution of the protein with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione, the GST tag was cleaved using Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) overnight at 4°C. The cleaved protein was diluted to reduce the concentration of NaCl to ~40 mM using buffer containing 50 mM Tris pH 7.6 and 5 mM DTT. The diluted protein was subjected to Source Q anion exchange chromatography using an AKTA FPLC system using buffer containing 50 mM Tris pH 7.6, 5 mM DTT with NaCl

gradient. Finally, the protein was applied to a Superdex-200 size-exclusion chromatography column using the AKTA FPLC system. The protein was concentrated and stored at -80°C in buffer containing 50 mM Tris pH 8.0, 150 mM NaCl and 5 mM DTT.

24 L of wild-type GST-Atg7^{NTD} was grown in BL21(DE3) Gold at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 23°C. The cells were spun down and resuspended in buffer containing 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM DTT and 2.5 mM PMSF. Cells were then lysed by a combination of sonication and the use of an EmulsiFlex-C5 cell homogenizer (Avestin). Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. Glutathione affinity chromatography was used as the first step in the protein purification protocol using 20 ml of glutathione sepharose beads. After passing the supernatant over the beads, the beads were washed with 300 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT. Following elution of the protein with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione, the GST tag was cleaved using Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) overnight at 4°C. The cleaved protein was diluted to reduce the concentration of NaCl to ~40 mM using buffer containing 50 mM MES pH 6.5 and 5 mM DTT. The diluted protein was subjected to gravity cation exchange using SP sepharose beads (Sigma) with buffer containing 50 mM MES pH 6.5, 5 mM DTT and 50 mM NaCl as wash buffer. The protein was eluted with buffer containing 50 mM MES pH 6.5, 5 mM DTT, 200 mM NaCl. Finally, the protein was applied to a Superdex-200 size-exclusion chromatography column using the AKTA fast protein liquid chromatography (FPLC) system. The protein was concentrated and stored at -80°C in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT.

Before the AUC experiment, all proteins were thawed and passed over a Superdex-200 size-exclusion chromatography column using the AKTA fast protein liquid chromatography (FPLC) system in a final buffer containing 50 mM Tris pH 7.6, 150 mM NaCl, 1 mM DTT.

Experiments were carried out in a ProteomeLab XL-I analytical ultracentrifuge with a four-hole or 8-hole (Beckman An-60Ti or An-50Ti rotor) and cells containing sapphire or quartz windows and charcoal-filled Epon double-sector centre pieces (Beckman Coulter, Fullerton, CA). The density and viscosity of the ultracentrifugation buffer, 50 mM Tris, 150 mM NaCl and 1 mM DTT at 4 and 20°C were calculated from its composition and the partial specific volume at 4 and 20°C and molecular weight of the proteins were calculated based on its amino acid composition using the software SEDNTERP [135]. Buffer from the size-exclusion column was used as the ultracentrifugation buffer and optical reference.

For the sedimentation velocity experiment the loading volume of 400 µl was identical for the reference and sample chambers of the double-sector centrepiece. Fringe displacement data at time intervals of 1.0 min were collected with the Rayleigh interference system for 10 hours at a rotor speed of 60,000 rpm and analysed with

SEDFIT software (www.analyticalultracentrifugation.com) using the model for continuous sedimentation coefficient distribution $c(s)$ with deconvolution of diffusional effects [136, 137].

The sedimentation coefficient distribution $c(s)$ was calculated with maximum entropy regularization at a confidence level of $p = 0.7$ and at a resolution of sedimentation coefficients of $n = 100$. The positions of the meniscus and bottom, as well as time-invariant and radial noises, were fitted. The s -value of the protein was determined by integration of the main peak of $c(s)$. The molecular weight distribution, $c(M)$, was calculated using the values determined from the $c(s)$ distribution. The best-fit weight-average frictional ratio (f/f_0) was determined from the $c(s)$ analysis, and a single value was assumed [137]. The s -values of the continuous sedimentation coefficient distribution, $c(s)$, were converted to standard condition (20 °C, water as solvent), $s_{20,w}$ -values, and plotted as shown.

Sedimentation equilibrium was attained at 24 h at a rotor temperature of 4 °C at increasing speeds of 12, 18, & 22 k rpm for Atg7 and Atg7^{CTD} and 18, 27 & 34 k rpm for Atg7^{NTD} [138]. Protein at concentrations of between 0.86 and 3.5 μ M (120 μ L) were loaded into double-sector centrepieces and absorbance distributions recorded at 280 nm in 0.001 cm radial intervals with 20 replicates for each point. Global least squares modelling were performed at multiple rotor speeds with the software SEDPHAT (www.analyticalultracentrifugation.com) using a reversible monomer-dimer self-association model [138].

Isothermal Titration Calorimetry (ITC)

Atg7 and Atg7^{NTD} were prepared as described in the AUC section above, with protein being finally stored in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT.

Atg3 was expressed in codon enhanced RIL cells as a N-terminal GST fusion. 24 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 16°C. The protein purification protocol was the same as for Atg7 shown in the AUC section above with the exception of the buffer used in Source Q anion exchange here was 50 mM Tris 7.6, 5 mM DTT with NaCl gradient.

His₆-Atg3^{ΔFR} was expressed in 12 L of RIL codon enhanced cells. Cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 16°C. The cells were spun down and resuspended in buffer containing 1X PBS, 20 mM Imidazole, 5 mM BME and 2.5 mM PMSF. Cells were then lysed by sonication. Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. The protein was first subject to nickel affinity chromatography by passing the supernatant over 10 ml of nickel sepharose beads. Beads were transferred to a column and washed with 150 ml of buffer containing 1X PBS, 20 mM Imidazole and 5 mM BME. The protein was then eluted with buffer containing 1X PBS, 250 mM Imidazole and 5 mM

BME. The His₆ tag was cleaved using TEV protease, overnight at 4°C. The protein was diluted to ~40 mM NaCl using buffer containing 25 mM Tris pH 8.0, 5 mM DTT. The diluted protein was then run over a Source Q anion exchange chromatography column using an AKTA FPLC system using buffer containing 25 mM Tris pH 8.0, 5 mM DTT with NaCl gradient. The protein was next applied to a Superdex-200 size-exclusion chromatography column using the AKTA FPLC system. The protein was stored at -80°C in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT.

Atg3^{FR_{pep}} was synthesized by the Hartwell Center for Bioinformatics and Biotechnology at St Jude Children's Research Hospital. Protein samples were buffer matched using NAP-5 desalting columns (GE) into buffer containing 25 mM Hepes pH 8.0, 150 mM NaCl and 2mM β-mercaptoethanol (BME). Peptide was weighed out and dissolved in the desalting buffer. Measurements were performed on a MicroCal ITC200. Atg7 or Atg7^{NTD} was placed in the sample cell at concentrations between 35-45 μM at 30 °C. Atg3 or Atg3^{FR_{pep}} was titrated in 2.0-2.5 μl injections with injection times of 4-5 sec. In the case of Atg3^{ΔFR}, Atg7 or Atg7^{NTD} were used at 94 μM in the sample cell due to the weaker binding of Atg3^{ΔFR}. The raw data was evaluated in Origin (V 7.0) to determine the values of the thermodynamic parameters.

Purification, Crystallization and Structure Determination for Atg7^{NTD}

Selenomethionine labeled Atg7^{NTD} was produced in BL21(DE3) Gold cells using autoinduction media. 12 L of cells were grown at 37°C till they were past the exponential phase and then turned the temperature down to 23°C, and let the cells go overnight. The protein was purified as described for Atg7^{NTD} the AUC section above. Following size-exclusion chromatography, the protein was concentrated and stored in buffer containing 25 mM Tri-Na-Citrate pH 5.6, 150 mM NaCl and 10mM DTT or 20 mM Tris pH 7.6, 150 mM NaCl and 10 mM DTT. Protein concentration used for crystallography varied between 20-60 mg/ml.

Screening was performed using a Mosquito crystallization robot (TTP Labtech) and manually, using drop ratio of 1:1 (200 nl protein: 200 nl well solution using the mosquito, and 1 μl protein: 1 μl well solution for manually set trays). All crystals were grown using the hanging drop vapor diffusion method against 100 μl of well solution. Initial crystals were plate clusters. Streak-seeding in 2 μl drops (1 μl protein: 1 μl well solution) was used to obtain single crystals. The crystals that yielded the structure grew at 4°C in 9.2 % Isopropanol, 0.1M potassium thiocyanate and 0.1M Citrate pH 4.0. The cryoprotectant condition used was 6 % glycerol, 6 % ethylene glycol and 6 % sucrose. Diffraction data was collected at APS, 24-ID-C. The data was indexed, integrated and scaled using HKL2000 [139]. PHENIX AutoSol was used for determining experimental phases and initial model building [121]. This was followed by repeated cycles of manual model building in Coot [140] and refinement with phenix.refine [121]. Details are available in Table 3-1.

Purification, Crystallization and Structure Determination for Atg7^{NTD}-Atg3^{FRpep}

Atg7^{NTD} was prepared as described in the AUC section above, with protein being finally stored in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT. Atg3^{FRpep} (synthesized by the Hartwell Center for Bioinformatics and Biotechnology at St. Jude) was dissolved in double-distilled water. Since the peptide had an overall negative charge, ammonium bicarbonate was added (to help dissolve the peptide) at a final concentration of 1.7 % (v/v), yielding a final peptide concentration of 46 mM. The peptide was mixed in at a eight- to nine-fold excess over protein with the final protein concentration being 30 mg/ml. Screening was carried out as described for selenomethionine labeled protein above using a mosquito crystallization robot (TTP Labtech) and manually. Crystals that yielded the structure grew at room temperature in 1.8 M ammonium sulfate, 0.1 M Bis-Tris pH 5.5 and 5 mM DTT. The cryoprotectant condition used was 12.5 % glycerol and 12.5 % ethylene glycol. Diffraction data was collected at ALS 822. The data was indexed, integrated and scaled using HKL2000 [139]. Molecular replacement was used to solve the structure of the Atg7^{NTD}-Atg3^{FRpep} complex using PHENIX AutoMR [121] with the Atg7^{NTD} structure (solved above) as the structural searchmodel. Repeated cycles of manual model building in Coot [140] and refinement with phenix.refine [121] generated the final structure. Details are available in Table 3-1.

Purification, Crystallization and Structure Determination for Atg7^{NTD}(P283D)

12 L of Atg7^{NTD}(P283D) was prepared as described in the AUC section above for Atg7^{NTD}, with protein being finally stored in buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 5 mM DTT.

Screening was carried out as described for selenomethionine labeled protein above using a mosquito crystallization robot (TTP Labtech) and manually. Crystals that yielded the structure grew at 4°C in 10% 2-Propanol, 0.05 M Citrate pH 4.4 and 0.1M potassium thiocyanate. The cryoprotectant condition used was 30% ethylene glycol. Diffraction data was collected at APS, 24-ID-C. The data was indexed, integrated and scaled using HKL2000 [139]. Molecular replacement was used to solve the structure of the Atg7^{NTD}(P283D) complex using PHENIX AutoMR with the Atg7^{NTD} structure (solved above) as the structural searchmodel. Repeated cycles of manual model building in Coot [140] and refinement with phenix.refine [121] generated the final structure. Details are available in Table 3-1.

Pull-down Assays

All Atg7^{NTD} constructs (ampicillin resistance) coded for a GST tag at their N-terminus that was cleavable with Thrombin protease. All Atg3 and Atg3^{FR} constructs (kanamycin resistance) contained a His₆-MBP tag at the N-terminus that was cleavable with TEV protease. The assay involved co-transforming an Atg7^{NTD} construct with an Atg3/Atg3^{FR} construct using combination of wild-type and mutant constructs into

BL21(DE3) Gold cells. For each co-transformation, 1 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 23°C. The cells were spun down and resuspended in buffer containing 1X PBS, 10 mM Imidazole, 5 mM BME and 2.5 mM PMSF. Cells were then lysed by sonication and the lysate was spun down to obtain a clarified supernatant that contained soluble protein. The supernatant for each co-transformation was split in half and subjected to glutathione affinity chromatography and nickel affinity chromatography by rocking the lysate with 150 µl of beads at 4°C for >1 hr, enabling the pull-down of both GST-Atg7^{NTD} and His₆-MBP-Atg3/Atg3^{FR} expressed in the same pool of cells so as to ensure expression of both proteins. Beads were then transferred to columns, where they were washed with 10 ml of buffer containing 1X PBS, 10 mM Imidazole, 5 mM BME. For the glutathione affinity chromatography, protein was eluted with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione whereas in the case of nickel affinity chromatography, protein was eluted with buffer containing 1X PBS, 250 mM Imidazole and 5 mM BME. Elutions were run on 15 % SDS-PAGE gels. In the event that protein bands for GST-Atg7^{NTD} and His₆-MBP-Atg3/Atg3^{FR} overlapped on the gel, cleavage with Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) enabled separation of the two bands (since only the tag for GST-Atg7^{NTD} was cleaved by Thrombin).

Generation of Model Showing Transfer in Trans

The sequence of Atg7 from residues 298-578 was submitted to the Phyre server [126] to obtain a model of Atg7^{CTD} minus Atg7^{C-ter} (Atg7^{CTDΔC-ter}). The topmost output model from Phyre was based on SCOP code: d1jw9b; i.d: 25%; E-value: 4.6e-28; estimated precision: 100%; Fold/PDB descriptor: Activating enzymes of the ubiquitin-like proteins; Superfamily: Activating enzymes of the ubiquitin-like proteins; Family: Molybdenum cofactor biosynthesis protein MoeB. This model was used to build part of the trans-model for transthiolation. First, the dimeric structure of MoeB (PDB code: 1JW9), generated using symmetry, was used as a template to align two copies of Atg7^{CTDΔC-ter} to build the dimeric model. This was done in Coot using the SSM Superpose function that performs structural alignments based on secondary structure [140]. Next, two Atg3 structures (PDB code: 2DYT) with residues 129-142 deleted, were then manually aligned with Atg7^{CTDΔC-ter} such that the Atg3 catalytic Cys was in close proximity to the catalytic Cys in Atg7^{CTDΔC-ter}. Finally, two Atg7^{NTD}-Atg3^{FRpep} structures were modeled within reasonable distances to their connections in the remainder of Atg7 and Atg3, respectively. Atg8 (PDB code: 2ZPN) was superimposed onto MoaD in Coot using SSM superpose [140]. The missing segments are shown as dashed lines and the ~60 C-terminal residues are represented as ovals (Atg7^{C-ter}).

Pulse-chase Transfer of [³²P]Atg8 from Atg7 to Atg3

Atg7 wild-type and mutant constructs for enzyme assays were expressed in BL21(DE3) Gold cells as N-terminal GST fusions. 6 -12 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 23°C. The cells were spun

down and resuspended in buffer containing 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM DTT and 2.5 mM PMSF. Cells were then lysed by sonication. Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. Protein was first subject to glutathione affinity chromatography by rocking with 500 μ l – 1 ml of glutathione sepharose beads, at 4°C for >1 hr. Beads were transferred to a column and washed with >40 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT. The protein was eluted with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione. The GST tag was cleaved using Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM), overnight at 4°C. The cleaved protein was then passed over a Superdex-200 size-exclusion chromatography column using an AKTA FPLC system. The protein was concentrated and stored in a final buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM DTT at a concentration minimally >1 mg/ml at -80°C.

Atg3 wild-type and mutant constructs for enzyme assays were expressed in codon enhanced RIL cells as N-terminal His₆-MBP fusions. 6 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 16°C. The cells were spun down and resuspended in buffer containing 1X PBS, 20 mM Imidazole, 5 mM BME and 2.5 mM PMSF. Cells were then lysed by sonication. Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. The protein was first subject to nickel affinity chromatography by rocking with 2 ml of nickel sepharose beads, at 4°C for >1 hr. Beads were transferred to a column and washed with >40 ml of buffer containing 1X PBS, 20 mM Imidazole and 5 mM BME. The protein was then eluted with buffer containing 1X PBS, 250 mM Imidazole and 5 mM BME. The His₆-MBP tag was cleaved using TEV protease, overnight at 4°C. The protein was diluted to ~40 mM NaCl using buffer containing 25 mM Tris pH 7.6, 1 mM DTT. The diluted protein was then run over a Source Q anion exchange chromatography column using an AKTA FPLC system using buffer containing 25 mM Tris pH 7.6, 1 mM DTT with NaCl gradient. The eluted protein was then desalted into buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM DTT and concentrated to ~30 mg/ml before storage at -80°C.

Atg7^{trans} protein was made by co-transforming a vector (ampicillin resistance) encoding GST-Atg7(P283D) with a second vector (kanamycin resistance) encoding His₆-Atg7(C507A) into BL21(DE3) Gold cells. 24 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 23°C. The cells were spun down and resuspended in buffer containing 1X PBS, 20 mM Imidazole, 5 mM BME and 2.5 mM PMSF. Cells were then lysed by sonication. Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. The protein was first subject to nickel affinity chromatography by stirring 5 ml of nickel sepharose beads with the supernatant. Beads were transferred to a column and washed with 150 ml of buffer containing 1X PBS, 20 mM Imidazole and 5 mM BME. The protein was then eluted with buffer containing 1X PBS, 250 mM Imidazole and 5 mM BME onto a column packed with 2 ml of glutathione sepharose beads. The glutathione sepharose beads were washed with >40 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT. Tagged Atg7^{trans} protein (homodimer composed of one GST-Atg7(P283D) monomer and one His₆-Atg7(C507A) monomer) was then eluted with buffer containing 50 mM Tris pH

8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione. The GST and His₆ tags were cleaved using TEV protease to cleave the His₆ tag and Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) to cleave the GST tag, overnight at 4°C. The cleaved protein was then passed over a Superdex-200 size-exclusion chromatography column using an AKTA FPLC system. The protein was concentrated and stored in a final buffer containing 20 mM Tris pH 7.6, 150 mM NaCl and 1 mM DTT at a concentration ~1 mg/ml at -80°C.

Atg7^{cis} protein was made by co-transforming a vector (ampicillin resistance) encoding GST-Atg7(P283D, C507A) with a second vector (kanamycin resistance) encoding His₆-Atg7 into codon enhanced Rosetta cells. Protein was expressed and purified the same as described above for Atg7^{trans}.

GST-Atg8 was expressed from the pGEX-2TK vector that yields protein with a N-terminal PKA tag for enabling labeling with a radio-labeled phosphate moiety. Atg8 was transformed into codon enhanced Rosetta cells and 6 L of cells were grown at 37°C to an OD₆₀₀ ~0.9 and induced with 0.6 mM IPTG overnight at 16°C. The cells were spun down and resuspended in buffer containing 50 mM Tris pH 7.6, 200 mM NaCl, 5 mM DTT and 2.5 mM PMSF. Cells were then lysed by sonication. Following this, the lysate was spun down to obtain a clarified supernatant that contained soluble protein. Protein was purified using glutathione affinity chromatography by passing supernatant over a column packed with 10 ml glutathione sepharose beads. After washing with 200 ml of buffer containing 50 mM Tris pH 7.6, 200 mM NaCl and 5 mM DTT, protein was eluted with buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 5 mM DTT and 10 mM glutathione. The GST tag was cleaved using Thrombin protease (with the addition of calcium chloride at a final concentration of 2.5 mM) overnight at 4°C. The cleaved protein was diluted to ~40mM NaCl using buffer containing 50 mM Hepes pH 7.0, 1 mM DTT. The diluted protein was then run over a Source S cation exchange chromatography column using an AKTA FPLC system using buffer containing 50 mM Hepes pH 7.0, 1 mM DTT with NaCl gradient. The protein was next passed over a Superdex-200 size-exclusion chromatography column using an AKTA FPLC system. The protein was stored in a final buffer containing 25 mM Tris pH 7.6, 150 mM NaCl and 1 mM DTT at a concentration of ~6 mg/ml at -80°C.

For the enzyme assays, Atg8 was labeled with ³²P at an N-terminal PKA site using PKA (New England Biolabs) and [γ -³²P]ATP as described previously for other UBLs [35]. Atg7 and mutant versions were then charged with [³²P]Atg8, giving rise to a pool of Atg7~[³²P]Atg8. This was done in a 60 μ l reaction volume with 50 mM HEPES pH 7.5, 2 mM ATP, 5 mM MgCl₂, 2 mg/ml BSA, 1 μ M Atg7 (wild-type or mutants) and 2.5 μ M [³²P]Atg8. This loading of Atg7 wild-type and mutants with [³²P]Atg8 was allowed to proceed for 7 min following which the reactions were desalted using Zeba spin columns (Thermo Scientific) into a buffer containing 50 mM HEPES pH 7.5 and 50 mM EDTA pH 8.0 to stop re-loading of Atg7. The transfer of [³²P]Atg8 from the Atg7 to Atg3 was carried out in 100 mM HEPES pH 7.5, 2 mg/ml BSA, 50 mM EDTA pH 8.0, 500 nM Atg3, with a 1:5 dilution of desalted Atg7~[³²P]Atg8. Since the Atg7^{trans} and Atg7^{cis} have half as many active sites and Atg3 binding sites as wild-type Atg7, twice as

much of these enzymes were used in the transfer reaction to Atg3 compared to wild-type Atg7. The reactions were quenched with 2X SDS loading buffer and run on Nu-PAGE 4-12% Bis-Tris gels in MOPS buffer (Invitrogen).

Cell-based Assays

All cells were cultured in DMEM supplemented with 10% FCS, 1mM pyruvate, 1mM glutamine, 100U per ml penicillin, 100ug per ml streptomycin. Cells were cultured in a humidified environment at 37°C with 5% CO₂. Phoenix Eco cells (provided by Gary Nolan, Stanford University) were transfected with LZRS Atg7 retroviral constructs using Lipofectamine 2000 (Invitrogen). Two-days post transfection, viral supernatant was harvested, filtered through a 0.2 μ m filter and polybrene was added to 5mg/ml final concentration. To generate Atg7 null murine embryonic fibroblasts (MEF) cell lines, timed matings of Atg7 +/- mice were setup [141]. At day E12, embryos were removed and primary MEFs were isolated. Atg7 null MEFs were identified by genotyping and transformed with SV40 plasmid (gift from Joseph Opferman, St. Jude Children's Research Hospital). SV40 transformed, Atg7 null MEF (1 x10⁵ cells per well of a 6 well dish) were infected with viral supernatant. Two days post-infection, transduced cells were selected using Zeocin (100mg/ml, Invitrogen). To ensure equal expression of the wild-type and mutant GFP-Atg7 fusion proteins, the stably transduced, polyclonal cell lines were sorted once by flow cytometry using an identical sort gate.

Cells were treated with Chloroquine (50 μ M) for varying times. Cells were then lysed on ice for 30 minutes in NP40 lysis buffer (1% NP40, 150mM NaCl, 1mM PMSF, Complete protease inhibitors (Roche)). Lysates were clarified by centrifugation at 10,000 g for 10 minutes at 4°C and assayed for total protein content using Biorad protein assay. DTT (50mM) and protein sample buffer (Biorad) were added, samples were boiled and separated on Criterion 4-12% SDS-PAGE gels (Biorad) (30mg of protein per lane) followed by transfer to nitrocellulose. Membranes were probed overnight at 4°C with primary antibodies against LC3 (1/1000 dilution, Cell Signaling, 3868S), Atg12 (1/1000 dilution, Cell Signaling, 2011), Atg7 (1/1000 dilution, Sigma, A2856) and actin (1/10,000 dilution, MP Biomedicals, 691001). Membranes were washed extensively, probed with the appropriate HRP-conjugated antibody and immunoreactive proteins were detected by Supersignal ECL (ThermoScientific).

CHAPTER 4. IMPLICATIONS AND FUTURE DIRECTIONS

Expansion on the E1 Theme

Eukaryotic UBLs and E1 appear to have originated from ancient prokaryotic biosynthetic pathways. *E. coli* MoeB and ThiS represent prokaryotic UBLs, which function with their respective E1 adenylation domains, MoeB and ThiF, in the biosynthesis of molybdopterin and thiamine respectively [3, 142]. Prior structural studies have shown that both monomers of the MoeB and ThiF adenylation domains are active and bind one molecule of MoeB and ThiS each, giving rise to a 2:2 complex of MoeB:MoeB and ThiF:ThiS [12, 13].

The Adenylation Domain and Beyond

The basic theme represented by MoeB and ThiF has been elaborated upon over the course of evolution, with related proteins performing a variety of functions. For example, the *E. coli* enzyme MccB sequence has two regions. The N-terminal ~90 amino acids are unique, and the C-terminal ~250 residues are homologous to MoeB/ThiF. Like MoeB and ThiF, MccB functions as a homodimer with two active units, with each active unit being composed of an MoeB/ThiF like adenylation domain, and a novel globular domain (Figure 4-1) [143]. Each globular domain is composed of regions from both protomers (Figure 4-1). However, MccB is unique among MoeB/ThiF homologs in that its substrate is not a UBL. Rather, MccB adenylates the MccA heptapeptide during the biosynthesis of the Trojan horse antibiotic microcin C7 (MccC7) [143, 144]. The MccA heptapeptide binds in a groove in MccB that has the adenylation domain on one side and the novel domain clamping down from the other side [143]. This demonstrates well how domains beyond the highly conserved adenylation domain contribute to performing specific functional tasks.

Eukaryotic E1s possess a core MoeB/ThiF-like adenylation domain and can have a variety of additions to this scaffold. Canonical E1s, such as those for ubiquitin, Nedd8 and Sumo, have an additional cysteine domain inserted in the adenylation domain and a C-terminal ubiquitin-fold domain (ufd) (Figure 4-2A) [23, 24, 101]. Notably, canonical E1s are pseudosymmetric in that they have two adenylation domains, but only one is active and able to bind a UBL [3]. Following activation of the UBL by the adenylation domain, the C-terminal tail of the UBL forms a thiolester bond with the catalytic cysteine housed within the cysteine domain [3]. In the final step, the canonical E1s recruit an E2 via their ufd domain in order for the transthiolation reaction to occur as shown by previous studies and our current study on the Sumo pathway [3, 145]. In this manner, canonical E1s utilize additional domains to catalyze transfer of their UBL to an E2.

Like MoeB and ThiF, the non-canonical E1 Atg7 functions as a homodimer [3, 77]. This E1 has two additional domains, per monomer, beyond the MoeB/ThiF-like adenylation domain scaffold: Atg7^{NTD} (Atg7 N-terminal domain, ~300 residues) and

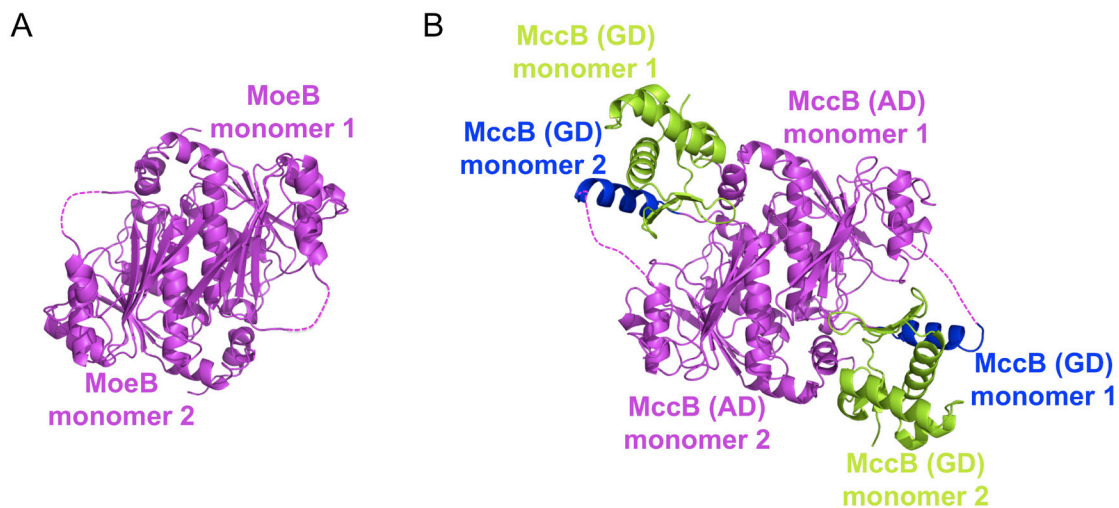


Figure 4-1. MccB is a prokaryotic protein with an adenylation domain at its core.

(A) Structure of homodimeric MoeB (magenta; PDB code: 1JW9) (B) The homodimeric MccB is composed of a MoeB/ThiF-like adenylation domain (AD, magenta) and a globular domain (GD, green and blue), providing two sites for catalysis. MccB AD from monomer 1 works together with its GD composed of parts from both monomers: the green region that is part of the opposite protomer (monomer 2) and the blue region that is part of the same polypeptide (monomer 1). The reverse is true for MccB AD from monomer 2. MccB structure's PDB code: 3H5N.

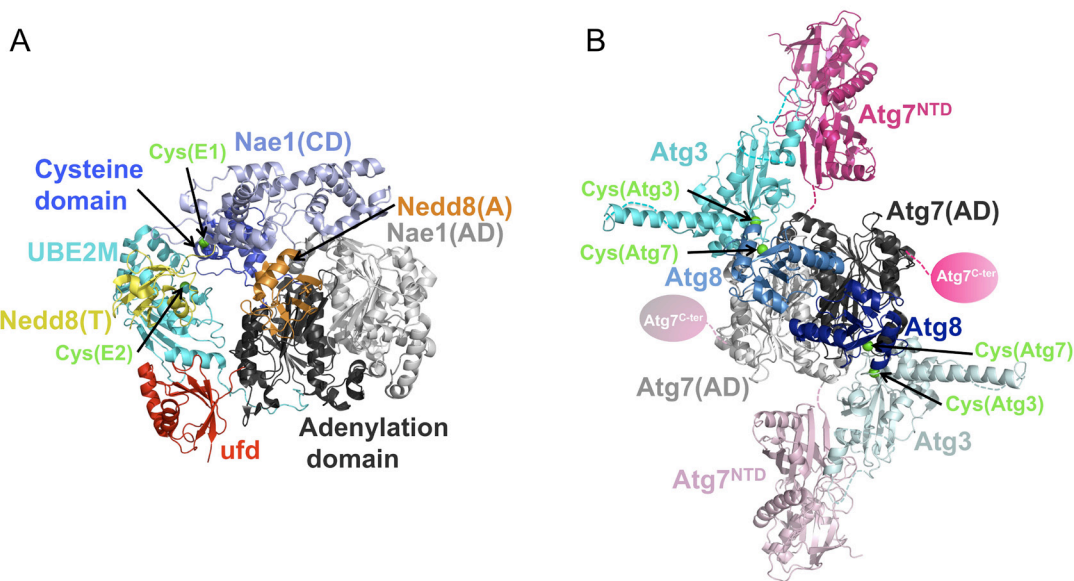


Figure 4-2. Beyond the MoeB/ThiF-like adenylation domain in eukaryotic E1s.

The adenylation domain dimer of both E1s is shown in light and dark grey. E2s are in shades of cyan. (A) Structure of E1 for Ned88 is composed of a heterodimer made up of Nae1 (AD : Adenylation Domain (light grey) and CD: Cysteine Domain (light blue)) and Uba3 (PDB code: 2NVU). Uba3 possess three domain crucial to E1 function: 1) an adenylation domain (magenta) that binds and activates Ned88; 2) a cysteine domain (blue) that houses the catalytic cysteine (green) that attacks the adenylylated Ned88 molecule; and 3) a ubiquitin-fold domain (ufd, red) that binds the E2 for transthiolation. Nae1 is the inactive monomer in this complex, unable to bind ATP or Sumo. (B) Model of Atg7 composed of a homodimer. Atg7, like canonical E1s, has domains to perform the three functions: 1) an adenylation domain (grey) that binds and activates Atg8; 2) a loop on the adenylation domain that houses the catalytic cysteine (green) that attacks the adenylylated Atg8 molecule; and 3) a unique N-terminal domain (Atg7^{NTD}, pink) that binds Atg3 for transthiolation. The function of Atg7^{C-ter} is unknown.

Atg7^{C-ter} (~60 C-terminal residues) (Figure 4-2B). On the basis of the work presented in Chapter 3, Atg7 now joins the list of E1s with an additional domain that has been structurally characterized. Atg7's N-terminal domain adopts a unique fold and mediates interactions with its E2, Atg3, in a unique manner, demonstrating how distinctive E1-E2 interactions can be incorporated beyond the adenylation domain scaffold (Figure 4-2B). Compared to the larger E1-E2 interface employed by canonical E1 ufd and E2s, Atg7 utilizes a small groove on its N-terminal domain for binding to its E2. Besides utilizing a small groove to bind Atg3, it is possible that Atg7^{NTD} has additional functions, either in Atg8 or Atg12 activation, or in other steps in the pathway that are yet to be determined. The ~60 residues at the C-terminus of Atg7 appear too short to adopt a ufd [64], but are required for soluble expression in *E. coli* and for various functions in yeast [77]. It remains to be seen if this domain has unique interactions with the UBLs, E2s or other macromolecules regulating autophagy.

It will be important to determine structures of accessory domains from other E1 enzymes and to understand how they contribute to enzyme functions. Two additional non-canonical E1s, Uba4 and Uba5, have additional regions outside the core MoeB/ThiF-like adenylation domain which are not well understood [3]. Uba4 and Uba5 are smaller than other E1 enzymes, possessing >400 residues per monomer compared, for instance, to the >600 residues per Atg7 monomer. Uba4 functions in sulfur transfer reactions and partners with its UBL, Urm1 (ubiquitin-related modifier 1). In addition to a short extension at the amino terminus of the adenylation domain, Uba4 possesses a rhodanase homology domain (RHD) at the carboxy terminus of the adenylation domain. It is the RHD domain is thought to catalyze sulfur transfer [3, 142]. Uba5 functions with its UBL, UFM1 (ubiquitin-fold modifier 1), whose biological function is currently unknown. The E2 enzyme for Ufm1 is UFC1 (UFM1-conjugating enzyme) whose only known E3 is UFL1 (UFM1-ligase) [3, 146, 147]. Uba5 is expressed as two isoforms, a full-length version and splice variant (1-404 and 57-404) [146, 148]. The role of the N-terminal residues is unknown, but these are not strongly conserved. These N-terminal residues are not required for activation of UFM1 either. Residues comprising the adenylation domain follow these N-terminal residues. A domain C-terminal to the adenylation domain is necessary for transfer of the UFM1 to UFC1, and likely serves to recruit UFC1, analogous to the ufd domains in canonical E1s [149]. Future studies on these enzymes will reveal their mechanism, shedding light on the variations on E1 themes.

Flexible Tethering of Domains for Functionality

Studying how various domains are built around the MoeB/ThiF-like adenylation scaffold reveals a trend among these enzymes: the additional domains are flexibly attached to the adenylation domain core. This flexibility allows for conformational changes that are necessary for these proteins to perform their functions. Here again the prokaryotic MccB serves as a starting point to demonstrate this trend. As mentioned above, MccB possess an adenylation domain and a globular domain, both of which serve to clamp its peptide substrate at its active site for catalysis. A study by Regni et al solved several structures of MccB in its free state and bound to various ligands [143].

Superimposition of these structures revealed that the adenylation and globular domains adopt a range of conformations with respect to one another, similar to the opening and closing of a clamshell.

This trend of flexibly tethered domains is also clearly observed in eukaryotic canonical E1s of the ubiquitin, Sumo and Nedd8 pathway. The cysteine domains of canonical E1s house the catalytic cysteine. This domain is attached into one adenylation domain at one point by a crossover loop and at another point by a re-entry loop. Based on previous structures, the distance between the catalytic cysteine and the adenylation active site was $\sim 30\text{\AA}$ [23, 24]. We now know, based on a structural study on the E1 for Sumo, that for thiolester formation, this domain rotates a 130° such that the catalytic cysteine can attack the C-terminus of the UBL (Figure 1-4) [27]. In a related vein, the ufd for the E1 for Nedd8, was shown to undergo a dramatic 120° rotation in order to bring the catalytic cysteine of the E2 in close proximity to the thiolester-bonded C-terminus of the UBL (Figures 1-6 and 4-2A) [34]. Based on the findings described herein, we observe the potential for flexibility between the Atg7^{NTD} and Atg7^{CTD}. Substitution of a conserved Pro to a Gly within the linker between the two domains had no effect on the overall transthiolation reaction. However, spacing between the two domains is important. Further studies directed towards understanding this flexibility-function relationship will be necessary. It also remains to be seen if other non-canonical E1s employ this scheme of flexible tethered domains for performing functional roles.

The Need for Homodimerization

An interesting feature of adenylation domains of MoeB and ThiF is that each active site contains residues from both monomers – highlighting the need for their homodimerization [12, 13, 19]. Such is also the case with MccB, where both active sites have residues from each monomer. A key question is whether the role of homodimerization is limited to active site complementation, or whether homodimerization may have additional functional consequences?

Our first insight into roles of homodimerization beyond active site complementation comes from MccB, where the globular domain that clamps down on the peptide substrate is composed of regions from both monomers (Figure 4-1). Deletion of the region from the opposite protomer completely eliminates activity [143]. Atg7 functions as a homodimer and our study reveals a novel role for homodimerization – transthiolation in trans – where the Atg3 bound to an Atg7^{NTD} on one monomer, accepts Atg8 bound to the catalytic cysteine from the opposite protomer (Figure 4-2B). It is likely that additional roles/implications of E1 homodimerization will be uncovered as we gain a better understanding of these enzymes.

The Remaining Mystery of Atg7 Function

Previous studies indicate that the majority of Atg7 in a reaction becomes thioesterified to Atg8 [150]. This implies that both monomers making up Atg7 become thioester linked to Atg8. Taken together with the transthiolation in-trans mechanism, we propose that two transthiolation reactions may occur in parallel to form two Atg3~Atg8 conjugates per Atg7 dimer. However, in general Atg7 activates and transfers two UBLs, Atg8 and Atg12, to their respective E2s, Atg3 and Atg10, which raises questions on the possible modes of UBL transfer used by Atg7: Can both UBLs be simultaneously activated and transferred? Is Atg12 transferred in cis or trans? Do Atg8 and Atg12 have any influence on each other's activation by Atg7 or transfer to their E2s? We therefore have many exciting questions awaiting answers. Difficulty in producing purified Atg12 precludes us from currently answering these questions. Future studies will be required to sort out the mechanism employed by Atg7 in transferring two UBLs to two E2s.

Therapeutic Targeting of the Autophagy Pathway

The Yin and Yang of Autophagy

Autophagy, under normal conditions, provides a housekeeping service by eliminating protein aggregates and damaged organelles to maintain cellular homeostasis, and additionally provide nutrients in periods of starvation. It even specifically targets mitochondria and microorganisms for lysosomal degradation [74-76]. Autophagy is traditionally thought of as a process that is cytoprotective and confers a survival advantage on cells. Downregulation/ disruption of this process has been implicated in several diseases and it is thought that the activation of autophagy in such cases would prove beneficial. On the flip side, inhibition of autophagy in some situations may actually prove beneficial [43]. This Yin-Yang situation in the context of autophagy will be discussed below keeping in mind that we are only just beginning to understand the complexity of this amazing process.

Studies have clearly shown a link between Atg genes and neurodegeneration. Komatsu et al. reported that knocking out Atg7, a gene essential for autophagy, in the nervous system of mice, showed behavioral defects, reduction in coordinated movement and death within 28 weeks of birth [151]. Furthermore, Atg7 loss caused massive neuronal loss in cerebral and cerebellar cortices. Importantly, these mice accumulated poly-ubiquitinated protein as inclusion bodies, a hallmark of neurological disease. A study at the same time by Hara et al. reported the onset of neurodegeneration in mice with Atg5 knocked out [152]. The mice in this study demonstrated increasing defects in motor function and showed the accumulation of inclusion bodies. The studies suggest that the constitutive clearance of unfavorable proteins via autophagy plays a central role in maintenance of healthy neurons.

Having established the role of autophagy proteins in preventing neurodegeneration, it is now interesting to consider the Yin-Yang of autophagy in

neuropathology using the example of Alzheimer's disease (AD) with the aim of reconciling whether autophagy should be activated or suppressed in the treatment of this disease. AD is a neurological disorder that is characterized by neural atrophy that is preceded by the formation of intraneuronal neurofibrillary tangles, composed of aggregated tau protein, and the deposition of β -amyloid ($A\beta$) outside the cell. It was found that expression of the Atg protein Beclin 1 was markedly reduced in the brains of AD patients to levels that would likely hamper autophagosome formation [153]. This proposition was further supported through genetic studies. When Beclin 1 levels were decreased in AD transgenic mice, they observed decreased neuronal autophagy and increased $A\beta$ deposition together with increased neurodegeneration. These mice showed decreased $A\beta$ accumulation upon increased expression of Beclin 1 [153, 154]. This supports the role for activation of autophagy as a means for countering neurodegeneration. Another aspect of AD pathology is the accumulation in large numbers of autophagosomes and other autophagic-vacuoles within dystrophic, swollen neurites of affected neurons, implicating autophagy in AD pathogenesis [155]. This accumulation of autophagosomes could be due to various reasons, such as a defect in autophagosome fusion with the lysosome and its subsequent degradation. It has been implied that the activation of autophagy in such situations of impaired autophagosome clearance worsens AD pathology and AD would benefit from therapeutic modulation of the process of autophagy [156]. Reconciliation for this conundrum is proposed in a review by Garcia-Arencibia et al [157]. They suggest that in the early stages, before the onset of gross pathology, autophagy may be defective due to the decreased levels of Beclin 1. As the AD pathology progresses, there may be other factors that prevent autophagosomes from being cleared, which in turn leads to accumulation of large numbers of autophagosome and other autophagic-vacuoles [157].

A second example that highlights the Yin-Yang of autophagy is cancer. It was found that the Atg gene Beclin 1 was monoallelically deleted in 40-75% of human breast, prostate and ovarian cancer [43]. Mice that are heterozygous for Beclin 1 have increased frequencies of spontaneous malignancies, development of hepatitis B virus-induced premalignant lesions and increased cellular proliferation with a concomitant decrease in autophagy [158, 159]. The knock out of Atg4C in mice also made them more susceptible to tumor formation [160]. Furthermore, Atg5 loss has shown to play a role in natural killer cell malignancies [161]. These examples establish the link between autophagy genes and cancer, suggesting that Atg genes are tumor suppressors. It is however presently unknown as to how the suppression of autophagy leads to cancer. One popular hypothesis for this is that autophagy is the guardian of the genome, as cells deficient in autophagy show a marked increase in DNA damage, which in turn allows for the accumulation of mutations and deterioration of genome integrity that then propels cancer initiation and progression [53, 162, 163]. Therefore, overall it appears that the presence of autophagy suppresses tumorigenesis, or does it? Tumor cells have the ability to activate autophagy in response to various stresses they may face to ensure their viability. This is especially true when they have defects in apoptosis that would normally eliminate tumor cells subject to deprivation of nutrients [55, 162, 164]. Under conditions of defective apoptosis, tumor cells subject to stress can survive for extended periods by relying on autophagy while, loss of Beclin 1 or Atg5 abrogates the survival ability [53]. It

is possible that autophagy prevents the initiation of cancer, but once tumors are established, autophagy would allow these cancer cells the ability to survive despite the environmental stresses.

This phenomenon of autophagy being utilized by cancer cells for their own benefit raises the interesting prospect of inhibiting autophagy in established cancers as a means of cancer therapy as suggested in a review by Mathew et al [54]. Their reasoning is that shutting off autophagy in cells resistant to apoptosis will kill them via acute necrotic cell death. One potential caveat to this is that death by necrosis leads to inflammation, which in turn increases tumor growth rate [43].

Autophagy can be considered a paradox in context of many diseases that will unravel itself with time. Such situations arise not only in cancer and neurodegeneration as mentioned herein, but in other areas such as infection and immunity, and liver diseases (reviewed in [43, 165, 166]). The need for autophagic modulation in disease states is a clear conclusion based on the above information.

Specific Targeting of Autophagy Proteins with Small Molecules

The process of autophagy is choreographed by a complex signaling network that is still being deciphered, making the modulation of autophagy a challenge. However, much work has been done towards obtaining small molecule inducers and inhibitors of autophagy and there are several reviews describing/discussing the current information [167, 168]. What is clear is that there are more inducers than inhibitors, with a significant subset of these small molecules, especially inhibitors, having no known mechanism of action. Most importantly, there are no treatments at present that specifically target autophagy proteins, preventing us from controlling autophagy in a precise manner. However, there may be precedent from a compound, MLN4924 (from Millennium Pharmaceuticals) that is being developed as a mechanistic inhibitor for the E1 for Nedd8 as an anticancer drug [169].

MLN4924 is a selective inhibitor of the E1 for Nedd8 that is structurally related to adenosine 5'-monophosphate (AMP) [169]. It is a potent inhibitor of the E1 for Nedd8 ($IC_{50}=0.004\mu M$) and also shows selectivity against E1s for ubiquitin, Sumo and Atg7 with IC_{50} s of 1.5, 8.2 and $>10\mu M$ respectively [169]. Inhibition of the E1 for Nedd8 prevents the activation of cullin-RING ligases, which in turn prevents the ubiquitination and subsequent degradation of cullin-RING substrates. These substrates play a role in cell-cycle progression and survival and continued inhibition of the Nedd8 pathway leads to apoptosis [169]. Notably, MLN4924 inhibits tumor xenograft growth in mouse studies [169]. Recently, the mechanism by which this inhibitor works has also been deciphered. MLN4924 forms a covalent adduct with Nedd8's C-terminus (Figure 4-3) with a K_d that is considerably less than 1nM [170]. Adduct formation requires a series of steps: First, the E1 binds ATP and Nedd8, and carries out the adenylation reaction for the activation of Nedd8's C-terminus. Next, Nedd8's C-terminus forms a thiolester bond with the E1's catalytic cysteine. A molecule of MLN4924 then binds the active site of the E1 (while in

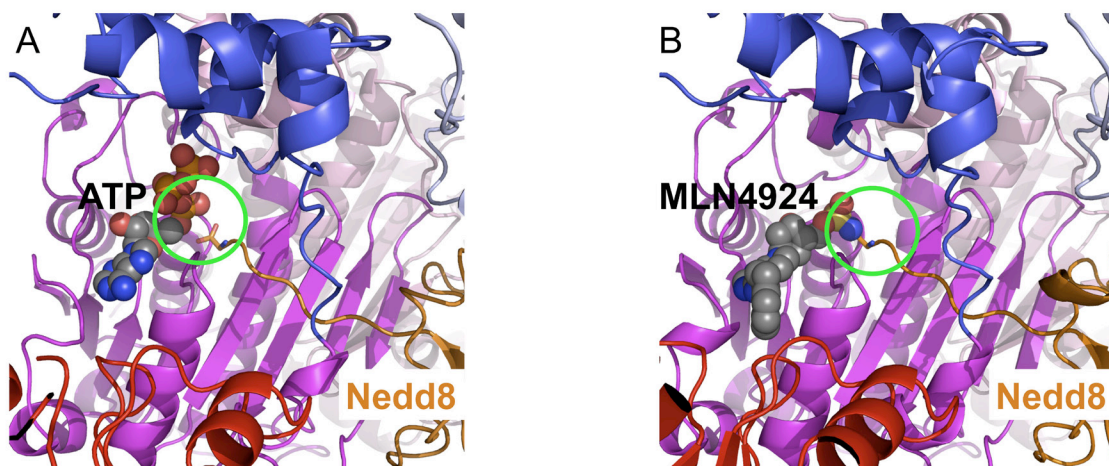


Figure 4-3. MLN4924 forms a covalent adduct with the C-terminus of Nedd8.

(A) ATP (spheres) and Nedd8's C-terminal tail in the E1's adenylation active site (highlighted with a green circle). The last residue, Gly76, of Nedd8 is shown in sticks. PDB code: 1R4N. (B) The same adenylation active site of the E1 for Nedd8 showing the covalent adduct (highlighted with a green circle) between MLN4924 (spheres) and Nedd8's C-terminus, with Gly76 shown in sticks. PDB code: 3GZN.

competition with ATP), where it acts as a nucleophile and attacks the thiolester bond between the C-terminus of Nedd8 and the catalytic cysteine of the E1, forming the final adduct that binds the E1 active site with high affinity (Figure 4-4) [170]. Conservation of E1 mechanism across various UBL pathways provides the potential to develop specific mechanistic inhibitors towards the rest of the E1s, including Atg7. Interestingly, compound 1, an analogue of MLN4924, was able to form adducts with ubiquitin, Nedd8, Sumo and Atg7 [170]. This is a potential avenue for specifically targeting Atg7 for the disruption of autophagy.

Our study now opens up an additional potential way for specifically targeting autophagy, namely through the specific inhibition of the Atg7^{NTD}-Atg3^{FR} interface. Targeting protein-protein interfaces through the use of small molecules poses many challenges (reviewed in [171]), a major hurdle being that protein-protein interfaces are typically extensive in the order of high hundreds to low thousands of Å² and do not typically possess small, deep cavities that may favor the binding of small molecules. However, there have been success stories, most notably, the development of a small molecule inhibitor of the p53-MDM2 interface (Figure 4-5A). This interface shares some similarities with our Atg7^{NTD}-Atg3^{FR} interface (Figure 4-5B). The p53-MDM2 inhibitor saga began with the determination of a crystal structure of the N-terminal domain of MDM2 in complex with a 15-residue peptide from p53 [172]. The structure revealed the p53 peptide adopting ~2 turns of a helix with the sidechains of three hydrophobic residues (Phe19, Trp23 and Leu26) from the peptide emanating into the MDM2 groove. The buried surface area is ~660 Å² and this interaction is considered a hot spot as majority of the binding affinity between p53 and MDM2 stems from this region, potentially allowing targeting by a small molecule [172, 173]. It took a considerable amount of time and effort, but there are now several classes of small molecules that can disrupt binding of p53 and MDM2 (reviewed in [173]). With the advancement of drug design, the Atg7^{NTD}-Atg3^{FR} interface may provide a potential drug target. Although the Atg7^{NTD}-Atg3^{FR} interface is slightly larger, 940 Å², and not quite as deep as the p53 binding groove on MDM2, it is still a hot spot that provides majority of the binding affinity between Atg3 and Atg7. Thus, it is possible that our studies of Atg7-Atg3 interactions may not only have revealed a novel mechanism for E1 to E2 UBL transfer, but could ultimately open the door to developing novel therapeutics.

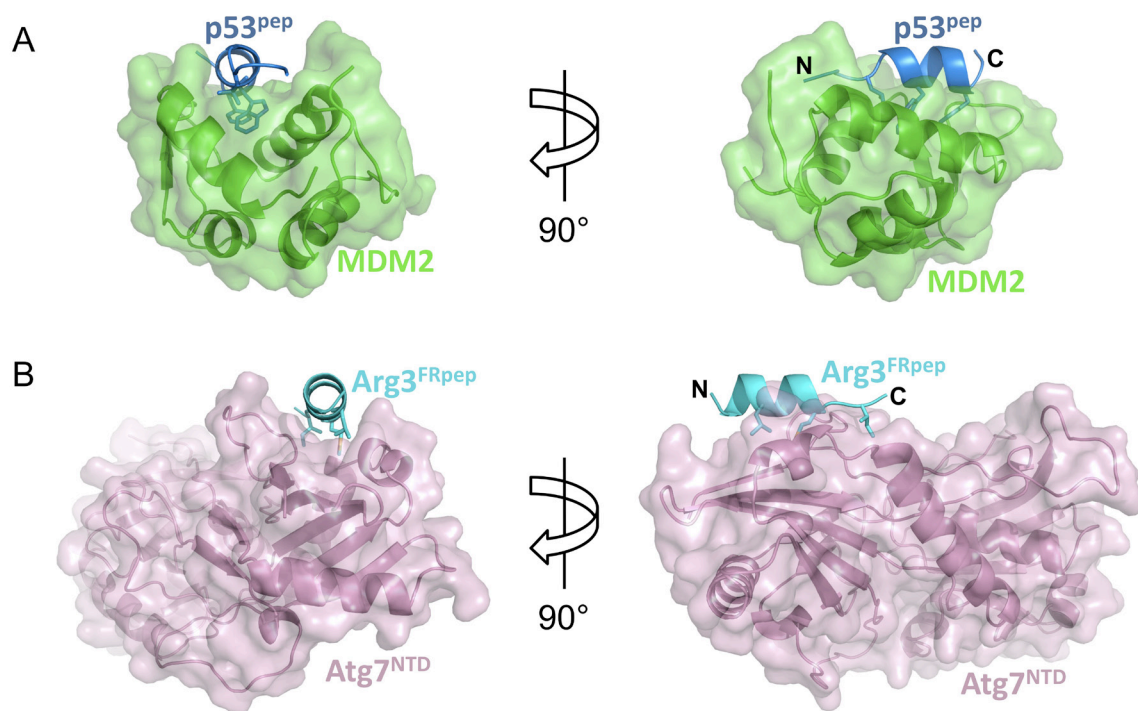


Figure 4-5. Comparison of the MDM2-p53^{pep} structure with the Atg7^{NTD}-Arg3^{FRpep} structure.

(A) Two views of the MDM2-p53^{pep} structure (MDM2 in green and p53^{pep} in blue; PDB code: 1YCR). Residues from the peptide involved in mediating key interactions are shown in sticks. The N and C terminus of peptide is labeled in the right panel. (B) Two views of the Atg7^{NTD}-Arg3^{FRpep} structure (Atg7^{NTD} in pink and Arg3^{FRpep} in cyan). Residues from the peptide involved in mediating key interactions are shown in sticks. The N and C terminus of peptide is labeled in the right panel.

LIST OF REFERENCES

1. Hochstrasser, M., *Origin and function of ubiquitin-like proteins*. Nature, 2009. **458**(7237): p. 422-9.
2. Weissman, A.M., *Themes and variations on ubiquitylation*. Nat Rev Mol Cell Biol, 2001. **2**(3): p. 169-78.
3. Schulman, B.A. and J.W. Harper, *Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways*. Nat Rev Mol Cell Biol, 2009. **10**(5): p. 319-31.
4. Vijay-Kumar, S., C.E. Bugg, and W.J. Cook, *Structure of ubiquitin refined at 1.8 Å resolution*. J Mol Biol, 1987. **194**(3): p. 531-44.
5. Bayer, P., et al., *Structure determination of the small ubiquitin-related modifier SUMO-1*. J Mol Biol, 1998. **280**(2): p. 275-86.
6. Whitby, F.G., et al., *Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes*. J Biol Chem, 1998. **273**(52): p. 34983-91.
7. Tang, Z., et al., *Protein interactions in the sumoylation cascade: lessons from X-ray structures*. FEBS J, 2008. **275**(12): p. 3003-15.
8. Hershko, A., A. Ciechanover, and I.A. Rose, *Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown*. J Biol Chem, 1981. **256**(4): p. 1525-8.
9. Ciechanover, A., et al., *"Covalent affinity" purification of ubiquitin-activating enzyme*. J Biol Chem, 1982. **257**(5): p. 2537-42.
10. Ciechanover, A., et al., *Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system*. Proc Natl Acad Sci U S A, 1981. **78**(2): p. 761-5.
11. Gareau, J.R. and C.D. Lima, *The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition*. Nat Rev Mol Cell Biol, 2010. **11**(12): p. 861-71.
12. Duda, D.M., et al., *Structural analysis of Escherichia coli ThiF*. J Mol Biol, 2005. **349**(4): p. 774-86.
13. Lake, M.W., et al., *Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex*. Nature, 2001. **414**(6861): p. 325-9.
14. Huang, D.T., et al., *Ubiquitin-like protein activation*. Oncogene, 2004. **23**(11): p. 1958-71.

15. Haas, A.L. and I.A. Rose, *The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis.* J Biol Chem, 1982. **257**(17): p. 10329-37.
16. Haas, A.L., et al., *Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation.* J Biol Chem, 1982. **257**(5): p. 2543-8.
17. Hershko, A., et al., *Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown.* J Biol Chem, 1983. **258**(13): p. 8206-14.
18. Haas, A.L. and P.M. Bright, *The resolution and characterization of putative ubiquitin carrier protein isozymes from rabbit reticulocytes.* J Biol Chem, 1988. **263**(26): p. 13258-67.
19. Lehmann, C., T.P. Begley, and S.E. Ealick, *Structure of the Escherichia coli ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis.* Biochemistry, 2006. **45**(1): p. 11-9.
20. Johnson, E.S., *Protein modification by SUMO.* Annu Rev Biochem, 2004. **73**: p. 355-82.
21. Guo, D., et al., *A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes.* Nat Genet, 2004. **36**(8): p. 837-41.
22. Owerbach, D., et al., *A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation.* Biochem Biophys Res Commun, 2005. **337**(2): p. 517-20.
23. Walden, H., M.S. Podgorski, and B.A. Schulman, *Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8.* Nature, 2003. **422**(6929): p. 330-4.
24. Lois, L.M. and C.D. Lima, *Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1.* EMBO J, 2005. **24**(3): p. 439-51.
25. Huang, D.T., et al., *Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1.* Mol Cell, 2005. **17**(3): p. 341-50.
26. Walden, H., et al., *The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1.* Mol Cell, 2003. **12**(6): p. 1427-37.
27. Olsen, S.K., et al., *Active site remodelling accompanies thioester bond formation in the SUMO E1.* Nature, 2010. **463**(7283): p. 906-12.
28. Bencsath, K.P., et al., *Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation.* J Biol Chem, 2002. **277**(49): p. 47938-45.

29. Wang, J., S. Cai, and Y. Chen, *Mechanism of E1-E2 interaction for the inhibition of Ubl adenylation*. J Biol Chem, 2010. **285**(43): p. 33457-62.
30. Sullivan, M.L. and R.D. Vierstra, *Cloning of a 16-kDa ubiquitin carrier protein from wheat and Arabidopsis thaliana. Identification of functional domains by in vitro mutagenesis*. J Biol Chem, 1991. **266**(35): p. 23878-85.
31. Pitluk, Z.W., et al., *Novel CDC34 (UBC3) ubiquitin-conjugating enzyme mutants obtained by charge-to-alanine scanning mutagenesis*. Mol Cell Biol, 1995. **15**(3): p. 1210-9.
32. Tong, H., et al., *Crystal structure of murine/human Ubc9 provides insight into the variability of the ubiquitin-conjugating system*. J Biol Chem, 1997. **272**(34): p. 21381-7.
33. Giraud, M.F., J.M. Desterro, and J.H. Naismith, *Structure of ubiquitin-conjugating enzyme 9 displays significant differences with other ubiquitin-conjugating enzymes which may reflect its specificity for sumo rather than ubiquitin*. Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 891-8.
34. Huang, D.T., et al., *Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity*. Nature, 2007. **445**(7126): p. 394-8.
35. Huang, D.T., et al., *E2-RING expansion of the NEDD8 cascade confers specificity to cullin modification*. Mol Cell, 2009. **33**(4): p. 483-95.
36. Bernier-Villamor, V., et al., *Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1*. Cell, 2002. **108**(3): p. 345-56.
37. Nakatogawa, H., et al., *Dynamics and diversity in autophagy mechanisms: lessons from yeast*. Nat Rev Mol Cell Biol, 2009. **10**(7): p. 458-67.
38. Mizushima, N. and B. Levine, *Autophagy in mammalian development and differentiation*. Nat Cell Biol, 2010. **12**(9): p. 823-30.
39. Yang, Z. and D.J. Klionsky, *An overview of the molecular mechanism of autophagy*. Curr Top Microbiol Immunol, 2009. **335**: p. 1-32.
40. Mizushima, N., *Autophagy: process and function*. Genes Dev, 2007. **21**(22): p. 2861-73.
41. Takeshige, K., et al., *Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction*. J Cell Biol, 1992. **119**(2): p. 301-11.
42. Mizushima, N., *Physiological functions of autophagy*. Curr Top Microbiol Immunol, 2009. **335**: p. 71-84.

43. Mizushima, N., et al., *Autophagy fights disease through cellular self-digestion*. Nature, 2008. **451**(7182): p. 1069-75.
44. Bernales, S., K.L. McDonald, and P. Walter, *Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response*. PLoS Biol, 2006. **4**(12): p. e423.
45. Yin, X.M., W.X. Ding, and W. Gao, *Autophagy in the liver*. Hepatology, 2008. **47**(5): p. 1773-85.
46. Yokota, S. and H. Dariush Fahimi, *Degradation of excess peroxisomes in mammalian liver cells by autophagy and other mechanisms*. Histochem Cell Biol, 2009. **131**(4): p. 455-8.
47. Mandl, J., et al., *Endoplasmic reticulum: nutrient sensor in physiology and pathology*. Trends Endocrinol Metab, 2009. **20**(4): p. 194-201.
48. Ruivo, R., et al., *Molecular and cellular basis of lysosomal transmembrane protein dysfunction*. Biochim Biophys Acta, 2009. **1793**(4): p. 636-49.
49. De Meyer, G.R. and W. Martinet, *Autophagy in the cardiovascular system*. Biochim Biophys Acta, 2009. **1793**(9): p. 1485-95.
50. Ballabio, A. and V. Gieselmann, *Lysosomal disorders: from storage to cellular damage*. Biochim Biophys Acta, 2009. **1793**(4): p. 684-96.
51. Chaturvedi, A. and S.K. Pierce, *Autophagy in immune cell regulation and dysregulation*. Curr Allergy Asthma Rep, 2009. **9**(5): p. 341-6.
52. Heath, R.J. and R.J. Xavier, *Autophagy, immunity and human disease*. Curr Opin Gastroenterol, 2009. **25**(6): p. 512-20.
53. White, E. and R.S. DiPaola, *The double-edged sword of autophagy modulation in cancer*. Clin Cancer Res, 2009. **15**(17): p. 5308-16.
54. Mathew, R., V. Karantza-Wadsworth, and E. White, *Role of autophagy in cancer*. Nat Rev Cancer, 2007. **7**(12): p. 961-7.
55. Degenhardt, K., et al., *Autophagy promotes tumor cell survival and restricts necrosis, inflammation, and tumorigenesis*. Cancer Cell, 2006. **10**(1): p. 51-64.
56. Ventruti, A. and A.M. Cuervo, *Autophagy and neurodegeneration*. Curr Neurol Neurosci Rep, 2007. **7**(5): p. 443-51.
57. Yu, W.H., et al., *Autophagic vacuoles are enriched in amyloid precursor protein-secretase activities: implications for beta-amyloid peptide over-production and localization in Alzheimer's disease*. Int J Biochem Cell Biol, 2004. **36**(12): p. 2531-40.

58. Wang, Y., et al., *Tau fragmentation, aggregation and clearance: the dual role of lysosomal processing*. Hum Mol Genet, 2009. **18**(21): p. 4153-70.
59. Cuervo, A.M., *Calorie restriction and aging: the ultimate "cleansing diet"*. J Gerontol A Biol Sci Med Sci, 2008. **63**(6): p. 547-9.
60. Cuervo, A.M., *Autophagy and aging: keeping that old broom working*. Trends Genet, 2008. **24**(12): p. 604-12.
61. Cuervo, A.M., et al., *Autophagy and aging: the importance of maintaining "clean" cells*. Autophagy, 2005. **1**(3): p. 131-40.
62. Kabeya, Y., et al., *Cis1/Atg31 is required for autophagosome formation in Saccharomyces cerevisiae*. Biochem Biophys Res Commun, 2007. **356**(2): p. 405-10.
63. Suzuki, K., et al., *Hierarchy of Atg proteins in pre-autophagosomal structure organization*. Genes Cells, 2007. **12**(2): p. 209-18.
64. Noda, N.N., Y. Ohsumi, and F. Inagaki, *ATG systems from the protein structural point of view*. Chem Rev, 2009. **109**(4): p. 1587-98.
65. Geng, J. and D.J. Klionsky, *The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. 'Protein modifications: beyond the usual suspects' review series*. EMBO Rep, 2008. **9**(9): p. 859-64.
66. Mizushima, N., et al., *A protein conjugation system essential for autophagy*. Nature, 1998. **395**(6700): p. 395-8.
67. Ichimura, Y., et al., *A ubiquitin-like system mediates protein lipidation*. Nature, 2000. **408**(6811): p. 488-92.
68. Xie, Z. and D.J. Klionsky, *Autophagosome formation: core machinery and adaptations*. Nat Cell Biol, 2007. **9**(10): p. 1102-9.
69. Sugawara, K., et al., *The crystal structure of microtubule-associated protein light chain 3, a mammalian homologue of Saccharomyces cerevisiae Atg8*. Genes Cells, 2004. **9**(7): p. 611-8.
70. Noda, N.N., et al., *Structural basis of target recognition by Atg8/LC3 during selective autophagy*. Genes Cells, 2008. **13**(12): p. 1211-8.
71. Nakatogawa, H., Y. Ichimura, and Y. Ohsumi, *Atg8, a ubiquitin-like protein required for autophagosome formation, mediates membrane tethering and hemifusion*. Cell, 2007. **130**(1): p. 165-78.

72. Klionsky, D.J., R. Cueva, and D.S. Yaver, *Aminopeptidase I of Saccharomyces cerevisiae is localized to the vacuole independent of the secretory pathway*. J Cell Biol, 1992. **119**(2): p. 287-99.
73. Yoshihisa, T. and Y. Anraku, *A novel pathway of import of alpha-mannosidase, a marker enzyme of vacuolar membrane, in Saccharomyces cerevisiae*. J Biol Chem, 1990. **265**(36): p. 22418-25.
74. Thurston, T.L., et al., *The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria*. Nat Immunol, 2009. **10**(11): p. 1215-21.
75. Kirkin, V., et al., *A role for ubiquitin in selective autophagy*. Mol Cell, 2009. **34**(3): p. 259-69.
76. Noda, N.N., Y. Ohsumi, and F. Inagaki, *Atg8-family interacting motif crucial for selective autophagy*. FEBS Lett, 2010. **584**(7): p. 1379-85.
77. Komatsu, M., et al., *The C-terminal region of an Apg7p/Cvt2p is required for homodimerization and is essential for its E1 activity and E1-E2 complex formation*. J Biol Chem, 2001. **276**(13): p. 9846-54.
78. Yamada, Y., et al., *The crystal structure of Atg3, an autophagy-related ubiquitin carrier protein (E2) enzyme that mediates Atg8 lipidation*. J Biol Chem, 2007. **282**(11): p. 8036-43.
79. Hanada, T., et al., *The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy*. J Biol Chem, 2007. **282**(52): p. 37298-302.
80. Geiss-Friedlander, R. and F. Melchior, *Concepts in sumoylation: a decade on*. Nat Rev Mol Cell Biol, 2007. **8**(12): p. 947-56.
81. Hannoun, Z., et al., *Post-translational modification by SUMO*. Toxicology, 2010. **278**(3): p. 288-93.
82. Ulrich, H.D., *The SUMO system: an overview*. Methods Mol Biol, 2009. **497**: p. 3-16.
83. Takahashi, Y., et al., *Smt3, a SUMO-1 homolog, is conjugated to Cdc3, a component of septin rings at the mother-bud neck in budding yeast*. Biochem Biophys Res Commun, 1999. **259**(3): p. 582-7.
84. Johnson, E.S. and G. Blobel, *Cell cycle-regulated attachment of the ubiquitin-related protein SUMO to the yeast septins*. J Cell Biol, 1999. **147**(5): p. 981-94.
85. Biggins, S., et al., *Genes involved in sister chromatid separation and segregation in the budding yeast Saccharomyces cerevisiae*. Genetics, 2001. **159**(2): p. 453-70.

86. Johnson, E.S., et al., *The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer*. EMBO J, 1997. **16**(18): p. 5509-19.
87. Gong, L., et al., *Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex*. FEBS Lett, 1999. **448**(1): p. 185-9.
88. Desterro, J.M., et al., *Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1*. J Biol Chem, 1999. **274**(15): p. 10618-24.
89. Okuma, T., et al., *In vitro SUMO-1 modification requires two enzymatic steps, E1 and E2*. Biochem Biophys Res Commun, 1999. **254**(3): p. 693-8.
90. Johnson, E.S. and G. Blobel, *Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p*. J Biol Chem, 1997. **272**(43): p. 26799-802.
91. Desterro, J.M., J. Thomson, and R.T. Hay, *Ubch9 conjugates SUMO but not ubiquitin*. FEBS Lett, 1997. **417**(3): p. 297-300.
92. Schwarz, S.E., et al., *The ubiquitin-like proteins SMT3 and SUMO-1 are conjugated by the UBC9 E2 enzyme*. Proc Natl Acad Sci U S A, 1998. **95**(2): p. 560-4.
93. Saitoh, H., et al., *Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2*. Curr Biol, 1998. **8**(2): p. 121-4.
94. Yunus, A.A. and C.D. Lima, *Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway*. Nat Struct Mol Biol, 2006. **13**(6): p. 491-9.
95. Ulrich, H.D., *The fast-growing business of SUMO chains*. Mol Cell, 2008. **32**(3): p. 301-5.
96. Lin, D., et al., *Identification of a substrate recognition site on Ubc9*. J Biol Chem, 2002. **277**(24): p. 21740-8.
97. Tatham, M.H., Y. Chen, and R.T. Hay, *Role of two residues proximal to the active site of Ubc9 in substrate recognition by the Ubc9.SUMO-1 thiolester complex*. Biochemistry, 2003. **42**(11): p. 3168-79.
98. Reverter, D. and C.D. Lima, *Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex*. Nature, 2005. **435**(7042): p. 687-92.
99. Yunus, A.A. and C.D. Lima, *Structure of the Siz/PIAS SUMO E3 ligase Siz1 and determinants required for SUMO modification of PCNA*. Mol Cell, 2009. **35**(5): p. 669-82.

100. Duan, X., et al., *Structural and functional insights into the roles of the Mms21 subunit of the Smc5/6 complex*. Mol Cell, 2009. **35**(5): p. 657-68.
101. Lee, I. and H. Schindelin, *Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes*. Cell, 2008. **134**(2): p. 268-78.
102. Wang, J., et al., *The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications*. Mol Cell, 2007. **27**(2): p. 228-37.
103. Jin, J., et al., *Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging*. Nature, 2007. **447**(7148): p. 1135-8.
104. Durfee, L.A., M.L. Kelley, and J.M. Huibregtse, *The basis for selective E1-E2 interactions in the ISG15 conjugation system*. J Biol Chem, 2008. **283**(35): p. 23895-902.
105. Wang, J., et al., *Conformational transition associated with E1-E2 interaction in small ubiquitin-like modifications*. J Biol Chem, 2009. **284**(30): p. 20340-8.
106. van Waardenburg, R.C., et al., *Distinct functional domains of Ubc9 dictate cell survival and resistance to genotoxic stress*. Mol Cell Biol, 2006. **26**(13): p. 4958-69.
107. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination*. Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 5): p. 905-21.
108. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D*. Nucleic Acids Res, 2010. **38**: p. W545-9.
109. Tatham, M.H., et al., *Unique binding interactions among Ubc9, SUMO and RanBP2 reveal a mechanism for SUMO paralog selection*. Nat Struct Mol Biol, 2005. **12**(1): p. 67-74.
110. Pichler, A., et al., *The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type*. Nat Struct Mol Biol, 2004. **11**(10): p. 984-91.
111. Johnson, E.S. and A.A. Gupta, *An E3-like factor that promotes SUMO conjugation to the yeast septins*. Cell, 2001. **106**(6): p. 735-44.
112. Zheng, N., et al., *Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases*. Cell, 2000. **102**(4): p. 533-9.
113. Duda, D.M., et al., *Structure of a SUMO-binding-motif mimic bound to Smt3p-Ubc9p: conservation of a non-covalent ubiquitin-like protein-E2 complex as a platform for selective interactions within a SUMO pathway*. J Mol Biol, 2007. **369**(3): p. 619-30.

114. Knipscheer, P., et al., *Ubc9 sumoylation regulates SUMO target discrimination*. Mol Cell, 2008. **31**(3): p. 371-82.
115. Knipscheer, P., et al., *Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation*. EMBO J, 2007. **26**(11): p. 2797-807.
116. Capili, A.D. and C.D. Lima, *Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction*. J Mol Biol, 2007. **369**(3): p. 608-18.
117. Otwinowski, Z.M., *Processing of X-ray Data Collected in Oscillation Mode*, in *Methods in Enzymology, Macromolecular Crystallography*. 1997. p. 307-326.
118. Storoni, L.C., A.J. McCoy, and R.J. Read, *Likelihood-enhanced fast rotation functions*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 3): p. 432-8.
119. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
120. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions*. Methods Enzymol, 2003. **374**: p. 300-21.
121. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
122. McEwan, D.G. and I. Dikic, *The Three Musketeers of Autophagy: phosphorylation, ubiquitylation and acetylation*. Trends Cell Biol, 2011.
123. Yorimitsu, T. and D.J. Klionsky, *Autophagy: molecular machinery for self-eating*. Cell Death Differ, 2005. **12 Suppl 2**: p. 1542-52.
124. Suzuki, K. and Y. Ohsumi, *Molecular machinery of autophagosome formation in yeast, Saccharomyces cerevisiae*. FEBS Lett, 2007. **581**(11): p. 2156-61.
125. Ichimura, Y., et al., *In vivo and in vitro reconstitution of Atg8 conjugation essential for autophagy*. J Biol Chem, 2004. **279**(39): p. 40584-92.
126. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server*. Nat Protoc, 2009. **4**(3): p. 363-71.
127. Tran, H.J., et al., *Structure of the Jab1/MPN domain and its implications for proteasome function*. Biochemistry, 2003. **42**(39): p. 11460-5.
128. Sanches, M., et al., *The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer*. J Mol Biol, 2007. **370**(5): p. 846-55.

129. Pena, V., et al., *Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa*. Mol Cell, 2007. **25**(4): p. 615-24.
130. Xie, K., et al., *The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1*. Proc Natl Acad Sci U S A, 2004. **101**(21): p. 8114-9.
131. Johansson, E., et al., *Structural, kinetic, and mutational studies of the zinc ion environment in tetrameric cytidine deaminase*. Biochemistry, 2004. **43**(20): p. 6020-9.
132. Sato, Y., et al., *Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains*. Nature, 2008. **455**(7211): p. 358-62.
133. Deprez, C., et al., *Solution structure of the E.coli TolA C-terminal domain reveals conformational changes upon binding to the phage g3p N-terminal domain*. J Mol Biol, 2005. **346**(4): p. 1047-57.
134. Werner, A.B., et al., *TRAIL receptor and CD95 signal to mitochondria via FADD, caspase-8/10, Bid, and Bax but differentially regulate events downstream from truncated Bid*. J Biol Chem, 2002. **277**(43): p. 40760-7.
135. Laue, T.M., Shah, B.D., Ridgeway, T.M., and Pelletier, S.L., *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, ed. S.E. Harding, Rowe, A.J. and Horton, J.C. 1992, Cambridge: The Royal Society of Chemistry.
136. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling*. Biophys J, 2000. **78**(3): p. 1606-19.
137. Schuck, P., et al., *Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems*. Biophys J, 2002. **82**(2): p. 1096-111.
138. Vistica, J., et al., *Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition*. Anal Biochem, 2004. **326**(2): p. 234-56.
139. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode.*, in *In Methods in Enzymology, Macromolecular Crystallography, part A*, C.W. Carter and R.M. Sweet, Editors. 1997. p. 307-326.
140. Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
141. Sanjuan, M.A., et al., *Toll-like receptor signalling in macrophages links the autophagy pathway to phagocytosis*. Nature, 2007. **450**(7173): p. 1253-7.

142. Hochstrasser, M., *Evolution and function of ubiquitin-like protein-conjugation systems*. Nat Cell Biol, 2000. **2**(8): p. E153-7.
143. Regni, C.A., et al., *How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic*. EMBO J, 2009. **28**(13): p. 1953-64.
144. Duquesne, S., et al., *Structural and functional diversity of microcins, gene-encoded antibacterial peptides from enterobacteria*. J Mol Microbiol Biotechnol, 2007. **13**(4): p. 200-9.
145. Wang, J., et al., *Crystal structure of UBA2(ufd)-Ubc9: insights into E1-E2 interactions in Sumo pathways*. PLoS One, 2010. **5**(12): p. e15805.
146. Komatsu, M., et al., *A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier*. EMBO J, 2004. **23**(9): p. 1977-86.
147. Tatsumi, K., et al., *A novel type of E3 ligase for the Ufm1 conjugation system*. J Biol Chem, 2010. **285**(8): p. 5417-27.
148. Zheng, M., et al., *UBE1DC1, an ubiquitin-activating enzyme, activates two different ubiquitin-like proteins*. J Cell Biochem, 2008. **104**(6): p. 2324-34.
149. Bacik, J.P., et al., *Crystal structure of the human ubiquitin-activating enzyme 5 (UBA5) bound to ATP: mechanistic insights into a minimalistic E1 enzyme*. J Biol Chem, 2010. **285**(26): p. 20273-80.
150. Oh-oka, K., H. Nakatogawa, and Y. Ohsumi, *Physiological pH and acidic phospholipids contribute to substrate specificity in lipidation of Atg8*. J Biol Chem, 2008. **283**(32): p. 21847-52.
151. Komatsu, M., et al., *Loss of autophagy in the central nervous system causes neurodegeneration in mice*. Nature, 2006. **441**(7095): p. 880-4.
152. Hara, T., et al., *Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice*. Nature, 2006. **441**(7095): p. 885-9.
153. Pickford, F., et al., *The autophagy-related protein beclin 1 shows reduced expression in early Alzheimer disease and regulates amyloid beta accumulation in mice*. J Clin Invest, 2008. **118**(6): p. 2190-9.
154. Spencer, B., et al., *Beclin 1 gene transfer activates autophagy and ameliorates the neurodegenerative pathology in alpha-synuclein models of Parkinson's and Lewy body diseases*. J Neurosci, 2009. **29**(43): p. 13578-88.
155. Nixon, R.A., et al., *Extensive involvement of autophagy in Alzheimer disease: an immuno-electron microscopy study*. J Neuropathol Exp Neurol, 2005. **64**(2): p. 113-22.

156. Boland, B., et al., *Autophagy induction and autophagosome clearance in neurons: relationship to autophagic pathology in Alzheimer's disease*. J Neurosci, 2008. **28**(27): p. 6926-37.
157. Garcia-Arencibia, M., et al., *Autophagy, a guardian against neurodegeneration*. Semin Cell Dev Biol, 2010. **21**(7): p. 691-8.
158. Qu, X., et al., *Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene*. J Clin Invest, 2003. **112**(12): p. 1809-20.
159. Yue, Z., et al., *Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor*. Proc Natl Acad Sci U S A, 2003. **100**(25): p. 15077-82.
160. Marino, G., et al., *Tissue-specific autophagy alterations and increased tumorigenesis in mice deficient in Atg4C/autophagin-3*. J Biol Chem, 2007. **282**(25): p. 18573-83.
161. Iqbal, J., et al., *Genomic analyses reveal global functional alterations that promote tumor growth and novel tumor suppressor genes in natural killer-cell malignancies*. Leukemia, 2009. **23**(6): p. 1139-51.
162. Karantza-Wadsworth, V., et al., *Autophagy mitigates metabolic stress and genome damage in mammary tumorigenesis*. Genes Dev, 2007. **21**(13): p. 1621-35.
163. Mathew, R., et al., *Autophagy suppresses tumor progression by limiting chromosomal instability*. Genes Dev, 2007. **21**(11): p. 1367-81.
164. Lum, J.J., et al., *Growth factor regulation of autophagy and cell survival in the absence of apoptosis*. Cell, 2005. **120**(2): p. 237-48.
165. Rautou, P.E., et al., *Autophagy in liver diseases*. J Hepatol, 2010. **53**(6): p. 1123-34.
166. Levine, B. and V. Deretic, *Unveiling the roles of autophagy in innate and adaptive immunity*. Nat Rev Immunol, 2007. **7**(10): p. 767-77.
167. Rubinsztein, D.C., et al., *Potential therapeutic applications of autophagy*. Nat Rev Drug Discov, 2007. **6**(4): p. 304-12.
168. Fleming, A., et al., *Chemical modulators of autophagy as biological probes and potential therapeutics*. Nat Chem Biol, 2011. **7**(1): p. 9-17.
169. Soucy, T.A., et al., *An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer*. Nature, 2009. **458**(7239): p. 732-6.

170. Brownell, J.E., et al., *Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: the NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ*. Mol Cell, 2010. **37**(1): p. 102-11.
171. Arkin, M.R. and J.A. Wells, *Small-molecule inhibitors of protein-protein interactions: progressing towards the dream*. Nat Rev Drug Discov, 2004. **3**(4): p. 301-17.
172. Kussie, P.H., et al., *Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain*. Science, 1996. **274**(5289): p. 948-53.
173. Domling, A., *Small molecular weight protein-protein interaction antagonists: an insurmountable challenge?* Curr Opin Chem Biol, 2008. **12**(3): p. 281-91.

VITA

Asad Taherbhoy was born in Singapore in 1978. He received his Bachelors of Science (hons) from the National University of Singapore (NUS) in 2005, following which, he joined the Integrated Program in Biomedical Sciences at the University of Tennessee Health Science Center for graduate studies. He carried out his dissertation research in the lab of Dr. Brenda Schulman at St Jude Children's Research Hospital and is expected to receive his Ph.D degree in May 2011.